

UNICAMP

UNIVERSIDADE ESTADUAL DE
CAMPINAS

Instituto de Matemática, Estatística e
Computação Científica

MANUEL JESÚS PEREIRA BARAHONA

Inferência Bayesiana para Processos Assimétricos Fracionários

Campinas

2018

Manuel Jesús Pereira Barahona

Inferência Bayesiana para Processos Assimétricos Fracionários

Tese apresentada ao Instituto de Matemática,
Estatística e Computação Científica da Uni-
versidade Estadual de Campinas como parte
dos requisitos exigidos para a obtenção do
título de Doutor em Estatística.

Orientadora: Laura Leticia Ramos Rifo

Este exemplar corresponde à versão
final da Tese defendida pelo aluno Ma-
nuel Jesús Pereira Barahona e orien-
tada pela Profa. Dra. Laura Leticia
Ramos Rifo.

Campinas
2018

Agência(s) de fomento e nº(s) de processo(s): Não se aplica.

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

P414i Pereira Barahona, Manuel Jesus, 1984-
Inferência bayesiana para processos assimétricos fracionários / Manuel Jesus Pereira Barahona. – Campinas, SP : [s.n.], 2018.

Orientador: Laura Leticia Ramos Rifo.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Inferência bayesiana. 2. Verossimilhança (Estatística). 3. Métodos MCMC. I. Rifo, Laura Leticia Ramos, 1970-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Inferencia bayesiana para procesos asimétricos fraccionarios

Palavras-chave em inglês:

Bayesian inference

Likelihood (Statistics)

MCMC methods (Statistics)

Área de concentração: Estatística

Titulação: Doutor em Estatística

Banca examinadora:

Laura Leticia Ramos Rifo [Orientador]

Jesus Enrique Garcia

Paulo Regis Caron Ruffino

Julio Michael Stern

Rafael Izbicki

Data de defesa: 04-07-2018

Programa de Pós-Graduação: Estatística

**Tese de Doutorado defendida em 04 de julho de 2018 e aprovada
pela banca examinadora composta pelos Profs. Drs.**

Prof(a). Dr(a). LAURA LETICIA RAMOS RIFO

Prof(a). Dr(a). JESUS ENRIQUE GARCIA

Prof(a). Dr(a). PAULO REGIS CARON RUFFINO

Prof(a). Dr(a). JULIO MICHAEL STERN

Prof(a). Dr(a). RAFAEL IZBICKI

As respectivas assinaturas dos membros encontram-se na Ata de defesa

Agradecimentos

Agradeço à minha família, principalmente Ana Barahona. Também quero agradecer à minha professora Laura Ramos Rifo, pela confiança que ela me deu e pelo seu empenho e dedicação. Agradeço também à agência de fomento, CAPES (intermitente 2013-2014) pela entrega de bolsa de estudo.

Resumo

Neste trabalho estamos interessados em dois tipos de processos. Primeiro, em processos onde uma barreira de semipermeabilidade divide o espaço de estados, produzindo uma difusão assimétrica em cada lado. E segundo, aqueles que apresentam uma distribuição marginal assimétrica e caudas pesadas. O interesse de trabalhar neste tipo de problemas surge no contexto de modelagem de dados temporais como, por exemplo, preços de ativos, dados climáticos, dados temporais bidimensionais, etc. Em particular, trabalhamos com métodos MCMC e Bayesianos livres de verossimilhança. Este último é aplicado a modelos onde a função de verossimilhança é analiticamente intratável ou computacionalmente proibitiva. Apresentamos inferências bayesianas sobre os parâmetros envolvidos nesses tipos de processos. Além disso, discutimos o comportamento assintótico de alguns estimadores bayesianos para os parâmetros. Validamos as inferências obtidas através de conjuntos de dados simulados e dados reais.

Palavras-chave: Inferencia bayesiana. processos assimétricos. livre de verossimilhança.

Abstract

In this work we have interest in two classes of processes. First, in processes where a semi-permeable barrier splits the state space, producing a skewed diffusion that can have different rates on each side. And second, those that have distributional asymmetry and heavy tails. The interest to work in this type of problems arises in the context of temporal data modeling for example; options prices, climate data, temporary two-dimensional data, etc. In particular, we work with MCMC and likelihood-free Bayesian methods. The latter is applied to models where the likelihood function is analytically intractable or computationally prohibitive. We present Bayesian inferences about the parameters involved in these classes of processes. In addition, we discuss the asymptotic behavior of some Bayesian estimators for the parameters. We validate the inferences through simulated datasets and real data.

Keywords: Bayesian inference. asymmetric processes. free of likelihood.

Lista de ilustrações

Figura 1 – Em vermelho o movimento Browniano padrão, e em preto o movimento Browniano com processo subordinado T	28
Figura 2 – À esquerda, incrementos do movimento Browniano com processo subordinado T , e à direita, os incrementos do movimento Browniano padrão.	28
Figura 3 – O movimento Browniano geométrico, considerando em preto $\mu = 1$, $\sigma = 0.5$ e em vermelho $\mu = 0.5$, $\sigma = 0.5$	30
Figura 4 – À esquerda, simulação de FATGBM com $\mu = 1$, $\theta = -1$ e $\sigma = 1$, e à direita, o log de FATGBM.	32
Figura 5 – A posteriori, dependendo do vetor de tolerâncias ϵ	49
Figura 6 – Kernel de transição $q(t, x, y \tau, \theta)$, com $t = 1$, $\tau = 1$, considerando $\theta = 0.8$ e: (a) $x = 0.1$, (b) $x = 1$, (c) $x = 2$, (d) $x = 3$	53
Figura 7 – Distribuição amostral da moda a posteriori dadas trajetórias com n passos, $n = 100, 1000, 10000$ assumindo $\theta = 0$	57
Figura 8 – Média a posteriori e quartis $q(0.025)$ y $q(0.975)$, para trajetórias com n passos, $n = 100, 1000, 10000$, assumindo $\theta = 0$	58
Figura 9 – Intervalo de amostragem para a média a posteriori de θ , com $\sigma^2 = 1$, considerando $\theta \in [-1, 1]$ e: (a) $n = 100$, (b) $n = 1000$, (c) $n = 10000$ passos. Para cada gráfico, a linha sólida representa o 95% dos valores centrais da média a posteriori, a linha pontilhada mostra o mesmo intervalo considerando apenas as trajetórias que possuem cinco cruces ou mais, e a linha segmentada, aquelas com dez cruces ou mais.	59
Figura 10 – Distribuição da amostra da moda a posteriori, dadas as trajetórias de n passos, $n = 100, 1000, 10000$ assumindo $\theta = .5$	60
Figura 11 – Boxplot da distribuição amostral da média a posteriori, considerando o número de cruzamentos com a barreira.	61
Figura 12 – (a) Simulação de uma trajetória do processo, com $n = 1000$ passos, para $\theta = 0.8$ e $\sigma^2 = 1$. Note que a trajetória tende a ser mais positiva que negativa. (b) Densidades a posteriori marginais e conjuntas para (θ, σ^2) , com distribuição a priori para θ : Beta(1,1) em azul, Beta(0.5,0.5) em rosa, Beta(5,1) em ciano, Beta(1,5) em verde. A linha vertical magenta representa a hipótese $\theta = 0$	62
Figura 13 – Localização da área de alimentação, Calbuco, Chile.	63

Figura 14 – (a) Latitude de uma trajetória de um leão-marinho na costa de frente a Calbuco, no sul do Chile. (b) Densidades a posteriori marginal e conjunta para o parâmetro de assimetria θ , com distribuição a priori: Beta(1,1) em azul , Beta(0.5,0.5) em rosa, Beta(5,1) em ciano, Beta(1,5) em verde. A linha vertical magenta representa a hipótese $\theta = 0$	64
Figura 15 – O kernel de transição, variando o parâmetro de assimetria θ , com $t = 1$ e $\sigma = 1$	67
Figura 16 – O kernel de transição, variando o parâmetro de escala σ , com $t = 1$ e $\theta = 0$	68
Figura 17 – Média a posteriori conjunta de (θ, σ) , dependendo do vetor de tolerância ϵ e assumindo $\theta = 0$ e $\sigma^2 = 1$	71
Figura 18 – Média a posteriori conjunta de (θ, σ) , dependendo do vetor de tolerância ϵ e assumindo $\theta = .5$ e $\sigma^2 = 1$	72
Figura 19 – Moda a posteriori conjunta de (θ, σ) , assumindo que $\epsilon = 0.1$ e todas as configurações de parâmetros.	73

Lista de tabelas

Tabela 1	– Simulações e níveis de tolerância.	47
Tabela 2	– Média e desvio padrão para distribuição a posteriori, dependendo das tolerâncias e assumindo que $\mu = 0$	48
Tabela 3	– Média e desvio padrão para moda e média a posteriori, assumindo que $\theta = 0$	59
Tabela 4	– Estimativas a posteriori para (θ, σ^2) , considerando diferentes distribuições a priori para θ , tomando uma trajetória simulada com valores nominais $\theta = 0.8$ e $\sigma^2 = 1$. As duas últimas colunas mostram a evidência em favor da hipótese $\theta = 0$, dadas por o e -valor e o fator de Bayes.	63
Tabela 5	– Estimativas a posteriori para (θ, σ^2) , considerando diferentes distribuições a priori para θ , para a trajetória SAL observada. As últimas três colunas mostram a evidência em favor da hipótese $\theta = 0$, dada pelo e -valor, o fator de Bayes e a probabilidade a posteriori de $\theta < 0$	65
Tabela 6	– Estatísticos usados para avaliar a suficiência.	68
Tabela 7	– Configuração de parâmetros do modelo.	68
Tabela 8	– Resultado de AB e MSE para os estimadores média e moda a posteriori.	69
Tabela 9	– Média e desvio padrão da amostra para a média a posteriori, dependendo do vetor de tolerâncias e assumindo que $\theta = -0.5$ e $\sigma^2 = 1$	69
Tabela 10	– Média e desvio padrão da amostra para a média a posteriori, dependendo do vetor de tolerâncias e assumindo que $\theta = 0$ e $\sigma^2 = 1$	70
Tabela 11	– Média e desvio padrão da amostra para a média a posteriori, dependendo do vetor de tolerâncias e assumindo que $\theta = 0$ e $\sigma^2 = 1$	70
Tabela 12	– Média e desvio padrão da amostra para a moda a posteriori, assumindo que $\epsilon = 0.1$ e todas as configurações de parâmetros.	70

Lista de Algoritmos

Algoritmo 1 – Computação Bayesiana Aproximada (ABC)	42
Algoritmo 2 – ABC-MCMC.	43
Algoritmo 3 – ABC-PRC	45

Sumário

	Introdução	14
I	ELEMENTOS FUNDAMENTAIS	16
1	PRELIMINARES	17
1.0.1	Probabilidade	17
1.0.2	Lei dos Grandes Números	19
1.0.3	Processos Estocásticos	20
1.0.3.1	Movimento Browniano	20
II	REFERENCIAIS TEÓRICAS	22
2	PROCESSOS SBM E FATGBM	23
2.1	Movimento Browniano Assimétrico	23
2.1.1	Densidade do sBm	25
2.2	Processo de Lévy	25
2.2.1	Processos Subordinados	26
2.2.1.0.1	Subordinados.	26
2.3	Modelo de Black-Scholes	29
2.3.1	Modelo	29
2.3.2	Distribuição tipo G	30
2.4	FATGBM	31
2.4.1	Distribuição Teórica do Modelo.	32
2.4.2	Distribuição dos Incrementos do Processo Subordinados	33
2.4.2.1	Subordinado com distribuição Gama Inversa	33
2.4.2.2	Subordinado com distribuição de Gama	34
2.4.2.3	Subordinado com distribuição Gaussiana Inversa	34
2.4.3	Estrutura de dependência do Modelo	34
III	INFERÊNCIA BAYESIANA	36
3	ELEMENTOS DE INFERÊNCIA BAYESIANA	37
3.1	Inferência Bayesiana	37
3.1.1	Teoria de Decisão	37
3.1.2	Estimativa por Regiões	39

3.1.3	Teste de Hipótese Bayesiana	40
3.1.3.1	Teste de Significância	41
3.1.3.2	Critério de Informação Bayesiana	42
3.2	Computação Bayesiana Aproximada	42
3.3	Suficiência	44
3.3.1	Entropia e Informação Mútua.	45
3.3.2	Desigualdade do Processamento de Dados e Estatísticos Suficientes	46
IV	RESULTADOS OBTIDOS	50
4	ESTIMADORES BAYESIANOS PARA O SBM	51
4.0.1	Resumo	51
4.0.2	Formulação do modelo	52
4.0.3	Tempo de Saída	53
4.0.4	Verossimilhança e configurações a priori	54
4.0.5	Considerações sobre o comportamento assintótico de a posteriori.	56
4.1	Computação para dados	60
4.1.1	Dados simulados	60
4.1.2	Real data	62
5	ESTIMADORES BAYESIANOS PARA EL FATGBM	66
5.1	Formulação do Modelo	66
5.1.1	Simulações	66
5.1.2	Configuração de distribuição a priori	67
5.1.3	Resultados	69
6	CONSIDERAÇÕES FINAIS	74
6.0.1	Resultados sBm	74
6.0.2	Resultados FATGBM	75
	REFERÊNCIAS	76
	Glossário	81
	ANEXOS	82
	ANEXO A – CÓDIGOS FONTE	83
A.0.1	Metropolis	83

A.0.2	ABC	84
-------	---------------	----

Introdução

Ao analisar conjuntos de dados temporários a partir de uma variável contínua, um processo de difusão pode ser apropriado sob certas condições na distribuição dos incrementos, e tem sido de interesse para vários estudos na última década. O interesse de trabalhar com esse tipo de modelo surge no contexto da modelagem de dados temporários, como, por exemplo, gestão de risco de seguro, (HURST; PLATEN, 1997; WAGNER, 2008), preços dos ativos, turbulências financeiras, (HEYDE; LEONENKO, 2005), em processos de ruído e mistura indeterminada (BARNDORFF-NIELSEN, 1977), em outros. Tais fenômenos são observados no tempo e tem a particularidade de que a distribuição marginal tem caudas pesadas.

Neste trabalho estamos interessados em processos onde uma barreira semi-permeável divide o espaço de estados, produzindo uma difusão assimétrica que pode ter taxas diferentes em cada lado (HARRISON; SHEPP, 1981a), este processo é chamado movimento Browniano assimétrico (sBm). Outra motivação deste projeto é trabalhar com o processo “Movimento Browniano geométrico com atividade do tempo Fractal ” (FATGBM) que apresenta, por um lado, assimetria na distribuição dos incrementos e por outro lado caudas pesadas (HEYDE, 1999; HEYDE; GAY, 2002; HEYDE et al., 2001; BARNDORFF-NIELSEN et al., 2002).

É discutido o comportamento assintótico de algumas inferências bayesianas para o processo sBm. O análise proposto é validada por meio de simulações. Até o momento, na literatura específica, o único resultado é o análise assintótico do estimador de máxima verossimilhança, sob a hipótese de assimetria nula, representando um movimento Browniano padrão (LEJAY; MORDECKI; TORRES, 2014). Além disso, algumas inferências bayesianas são discutidas em um tipo de processo gerado a partir de uma generalização do movimento browniano e de processos de Lévy.

A estimação dos parâmetros da distribuição de probabilidade marginal usando técnicas de verossimilhança livre, permite aproximar a distribuição a posteriori de cada um dos parâmetros com base nos dados observados e nas funções de distribuição a priori. Tais técnicas são aplicadas em contextos onde as funções de verossimilhança são analiticamente inviáveis ou computacionalmente complexas para avaliar. Estes métodos tentam de evitar a avaliação explícita da verossimilhança, ver por exemplo (BEAUMONT; ZHANG; BALDING, 2002) e (RATMANN et al., 2009). Existem diferentes propostas destes métodos entre as mais importantes: ABC (Cálculo Bayesiano Aproximado), ABC-MCMC proposto por (MARJORAM et al., 2003) e ABC-PRC proposto por (SISSON; FAN; TANAKA, 2007).

A primeira parte do trabalho consistiu na definição dos processos estocásticos em questão e nas propriedades associadas a eles. Esses resultados permitiram estabelecer um modelo probabilístico e uma família paramétrica bem definida para os problemas abordados.

A segunda parte do trabalho tem como objetivo inferências para as famílias anteriores, em particular, utilizando métodos de inferência bayesiana e verossimilhança livre para a construção de estimadores e testes de hipóteses. Estas ferramentas foram validadas através de simulações e aplicações para conjuntos de dados reais.

Organização do trabalho

O documento está organizado da seguinte forma. O capítulo 1 apresenta elementos matemáticos fundamentais. O capítulo 2 fornece a configuração dos modelos, propriedades e as definições básicas. O capítulo 3 mostra elementos fundamentais da inferência bayesiana, bem como algoritmos livres verossimilhança. O capítulo 4 apresenta uma análise baseada em simulações sobre o comportamento assintótico dos estimadores bayesianos propostos para o movimento browniano assimétrico, além de aproximações numéricas, análise de dados reais e discussão sobre alguns aspectos dos resultados obtidos. Finalmente, o capítulo 5 apresenta os resultados obtidos ao estimar os parâmetros do processo FATGBM por meio de métodos de verossimilhança livre.

Parte I

Elementos fundamentais

1 Preliminares

Neste capítulo, apresenta fundamentos matemáticos relevantes, que permitem a construção de processos e análise de propriedades expostas no seguimento do trabalho. Para um tratamento completo dos tópicos expostos, veja por exemplo (JACOD; PROTTER, 2004; MÖRTERS; PERES, 2010; DURRETT, 2010; RESNICK, 2013)

1.0.1 Probabilidade

Uma σ -álgebra \mathcal{F} é uma classe não vazia de subconjuntos de Ω fechado para complementação, uniões e intersecções contáveis. Os postulados mínimos para \mathcal{F} seja uma σ -álgebra são:

1. $\phi \in \mathcal{F}$.
2. Se $A^c \in \mathcal{F}$, então $A \in \mathcal{F}$.
3. Se $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{F}$, então $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}$.

O par (Ω, \mathcal{F}) é chamado *espaço mensurável* e cada $A \in \mathcal{F}$ é chamado *conjunto mensurável*. Uma σ -álgebra de particular interesse é a σ -álgebra de Borel de um conjunto Ω , a qual é o menor das σ -álgebra que contém a topologia de Ω e é denotada por $\mathcal{B}(\Omega)$. Se considerarmos (Ω, \mathcal{F}) um espaço mensurável. Diremos que $\mu : \mathcal{F} \longrightarrow \overline{\mathbb{R}}$, é uma medida sobre (Ω, \mathcal{F}) se verifica:

Se $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{F}$ tal que $A_i \cap A_j = \phi$ com $i \neq j$, então,

$$\mu \left(\bigcup_{i \in \mathbb{N}} A_i \right) = \sum_{i \in \mathbb{N}} \mu(A_i).$$

Em particular,

- Se μ é uma medida sobre (Ω, \mathcal{F}) e também $\mu(\Omega) = 1$, então é dito que é uma *medida de probabilidade* sobre (Ω, \mathcal{F}) e é denotada por \mathbb{P} .
- Por outro lado, seja $A = (a_1, b_1) \times (a_2, b_2) \times \dots \times (a_n, b_n) \subset \mathbb{R}$, onde para cada $i = 1, \dots, n$, $-\infty < a_i < b_i < +\infty$, diremos que a medida μ definida sobre $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ é a *medida de Lebesgue* se

$$\mu(A) = \prod_{i=1}^n (b_i - a_i).$$

Se \mathbb{P} é uma medida de probabilidade sobre (Ω, \mathcal{F}) , diremos que $(\Omega, \mathcal{F}, \mathbb{P})$ é um *espaço de probabilidade* e os conjuntos A que pertencem à σ -álgebra \mathcal{F} são chamados de eventos. Suponhamos que $(\Omega, \mathcal{F}, \mathbb{P})$ é um espaço de probabilidade e (Ω', \mathcal{F}') é um espaço mensurável. Diremos que a aplicação $X : \Omega \rightarrow \Omega'$ é uma variável aleatória se verifica $\{X \in A\} \in \mathcal{F}$ para todo $A' \in \mathcal{F}'$, isto é, se X é $\mathcal{F} - \mathcal{F}'$ mensurável.

Se considerarmos X uma variável aleatória de $(\Omega, \mathcal{F}, \mathbb{P})$ sobre (Ω', \mathcal{F}') , é possível induzir uma medida de probabilidade em (Ω', \mathcal{F}') , se considerarmos a seguinte relação

$$\mathbb{P}\{X \in A\} = \mathbb{P}_X(A) \quad A \in \mathcal{F}'.$$

Esta medida é chamada *distribuição* da variável aleatória X e

$$\mathbb{E}(X) = \int_{\Omega} X d\mathbb{P} = \int_{\Omega'} x d\mathbb{P}_X,$$

é a *esperança* de X . Sejam X, Y variáveis aleatórias em um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ diremos que são *identicamente distribuídas* se para cada evento A na σ -álgebra \mathcal{F} verifica

$$\mathbb{P}_X(A) = \mathbb{P}_Y(A)$$

e é denotado por $X \stackrel{d}{=} Y$. Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade. Uma medida μ sobre (Ω, \mathcal{F}) se diz que é *absolutamente contínua* em relação a \mathbb{P} se para todo conjunto mensurável $A \in \mathcal{F}$ tal que $\mathbb{P}(A) = 0$, implica $\mu(A) = 0$. Diremos que μ é uma medida σ -finita sobre (Ω, \mathcal{F}) , se existe uma sucessão $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{F}$ tal que $\Omega = \cup_{i \in \mathbb{N}} A_i$ e além disso $\mu(A_i) < \infty$, para cada $n \in \mathbb{N}$. O próximo resultado é conhecido como o teorema de Radon-Nicodym e a demonstração pode ser encontrada em (HALMOS, 2017).

Teorema 1. *Suponhamos μ uma medida σ -finita, tal que μ é absolutamente contínua em relação a \mathbb{P} . Então existe uma função mensurável $g : \Omega \rightarrow \mathbb{R}^+$, tal que para tudo $A \in \mathcal{F}$ verifica-se*

$$\mathbb{P}(A) = \int_A g d\mu.$$

Esta função g é comumente denotado por $\frac{d\mathbb{P}}{d\mu}$ e é chamado *derivada de Radon-Nikodym* de \mathbb{P} com respeito a μ . Se \mathbb{P}_X é a distribuição da variável aleatória X sobre um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ e μ uma medida σ -finita sobre (Ω, \mathcal{F}) , define-se a *densidade* de X por:

$$p_X = \frac{d\mathbb{P}_X}{d\mu}.$$

Em particular, se a distribuição de uma variável aleatória X é absolutamente contínua em relação à medida de Lebesgue, então,

$$p_X = \frac{d\mathbb{P}_X}{dx},$$

e verifica

$$\mathbb{P}(X \leq x) = \mathbb{P}_X((-\infty, x]) = \int_{-\infty}^x p_X(t) dt.$$

A relação acima, é chamada a *Função de Distribuição Cumulativa*. Todas as distribuições vistas neste trabalho são absolutamente contínuas em relação à medida de Lebesgue.

1.0.2 Lei dos Grandes Números

Seja $\{X_i\}_{i=1, \dots, n}$ uma sucessão de variáveis aleatórias definidas no mesmo espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, diz-se que convergem *quase certamente* para a variável aleatória X se

$$\mathbb{P}\left(\left\{\omega \in \Omega : X_n(\omega) \xrightarrow{n \rightarrow \infty} X(\omega)\right\}\right) = 1,$$

e é denotado por, $X_n \xrightarrow{c.s.} X$.

Um teorema que determina a convergência quase-segura de

$$\frac{S_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i,$$

é a *lei forte dos grandes números*. La demostración puede ser encontrada en ([JACOD; PROTTER, 2004](#)).

Teorema 2. *Seja $\{X_n\}_{n \geq 1}$ uma sucessão de variáveis aleatórias independentes, distribuídas identicamente e definidas no mesmo espaço de probabilidade. Então,*

$$\frac{S_n}{n} \xrightarrow{c.s.} \mu, \tag{1.1}$$

se e somente se

$$\mathbb{E}(X_j) = \mu < \infty,$$

para tudo j .

O teorema 2 permite o método não-determinístico de aproximação de integrais conhecido como *Método de Monte Carlo*, que permite aproximar a integral

$$I = \int_A h(x) dx, \tag{1.2}$$

utilizando simulações x_1, x_2, \dots, x_N provenientes de uma conveniente função de densidade p_X , então aproximar I por meio de

$$\hat{I} = \frac{1}{N} \sum_{i=1}^N \frac{h(x_i)}{p_X(x_i)},$$

que, pela lei dos grandes números

$$\frac{1}{N} \sum_{i=1}^N \frac{h(x_i)}{p_X(x_i)} \xrightarrow{c.s.} \mathbb{E}_{p_X} \left(\frac{h(x)}{p_X(X)} \right) = I.$$

O método de Monte Carlo, ao contrário dos métodos determinísticos, é útil quando a dimensão de A em (1.2) é grande.

1.0.3 Processos Estocásticos

Na teoria da probabilidade, um processo estocástico é uma sucessão de variáveis aleatórias que evoluem de acordo com outra variável, geralmente o tempo. Cada uma das variáveis aleatórias no processo tem sua própria função de distribuição de probabilidade e pode ou não estar correlacionada entre elas.

Seja $T \neq \emptyset$ um conjunto de índices, diremos que a sucessão $\{X_t\}_{t \in T}$ é um *processo estocástico* sobre o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ se X_t é uma variável aleatória para cada $t \in T$.

Se considerarmos (T, \mathcal{F}_T) um espaço mensurável. À função

$$t \rightarrow X_t,$$

é chamada de *trajetória* do processo estocástico $\{X_t\}_{t \in T}$. Como $\{X_t : t \in T\}$ é um processo aleatório com espaço de estado $S \subset \mathbb{R}$ e o conjunto de índices T é o conjunto \mathbb{N} ou $[0, \infty)$, em cujo caso T é interpretado como espaço de tempo discreto e tempo contínuo, respectivamente. No que se segue, vamos considerar $T \subseteq \mathbb{R}$.

Definição 1. *Seja o processo estocástico $\{X_t\}_{t \in T}$, então*

- (i) *X tem **incrementos estacionários** se para $s, t \in T$ com $s \leq t$, o incremento $X_t - X_s$ tem a mesma distribuição que $X_{(t-s)}$.*
- (ii) *X tem **incrementos independentes** se para $t_0, t_1, t_2, \dots, t_n \in T$ com $t_0 < t_1 < t_2 < \dots < t_n$, os incrementos $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ são independentes.*

Processos estocásticos com incrementos estacionários e independentes são o coração do conceito de infinita divisibilidade que é análogo ao passeio aleatório que começa no tempo zero,

$$X_t = \int_0^t dY_t.$$

O processo $X(t)$ é infinitamente divisível se satisfiz a definição 1 e adicionalmente satisfaz a condição de continuidade em probabilidade e inicialização zero, isto é,

$$X_{s+t} \xrightarrow{\mathbb{P}} X_s, \text{ quando } t \rightarrow 0, X_0 = 0, \forall t, s \geq 0.$$

1.0.3.1 Movimento Browniano

Diremos que o processo estocástico $\{B_t\}_{t \geq 0}$ sobre o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ é um *Movimento Browniano* se verifica:

1. $B_0 = x$ quase certamente, isto é, $\mathbb{P}(B_0 = x) = 1$, com $x \in \mathbb{R}$;

2. $\{B_t\}_{t>0}$ tem incrementos independentes;
3. Para tudo $t \geq 0$ e $h > 0$, $B_{t+h} - B_t \sim N(0, h)$;
4. $t \rightarrow B_t$ é contínua quase-certeza.

Se $x = 0$, diz-se que $\{B_t\}_{t>0}$ é um *Movimento Browniano padrão*. Algumas propriedades do processo são:

1. A esperança do processo é:

$$\mathbb{E}(B_t) = 0. \quad (1.3)$$

2. A variância, para o movimento browniano, em cada tempo t , é

$$\text{Var}(B_t) = t. \quad (1.4)$$

3. A *função característica* definida, para $k \in \mathbb{R}$, por

$$\varphi_X(k) = \mathbb{E}(e^{ikX}),$$

para o movimento Browniano é,

$$\varphi_{B_t}(k) = e^{-\frac{t^2 k^2}{2}},$$

para cada $t > 0$.

4. A *covariância* entre duas variáveis aleatórias X e Y que é definida por

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y),$$

para o movimento Browniano se tem

$$\text{Cov}(B_t, B_s) = \min\{s, t\}.$$

5. Finalmente, a *Correlação* entre duas variáveis aleatórias X e Y , que é definida por

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}},$$

para o movimento Browniano é,

$$\text{Corr}(B_s, B_t) = \sqrt{\frac{s}{t}},$$

para $s < t$.

Em particular, para um tratamento mais completo do movimento Browniano, ver por exemplo (MÖRTERS; PERES, 2010).

Parte II

Referenciais teóricos

2 Processos sBm e FATGBM

2.1 Movimento Browniano Assimétrico

O interesse de trabalhar com o movimento Browniano assimétrico (sBm) se origina em sua relação com processos de difusão com descontinuidade e problemas de barreiras semi-permeáveis, entre outros. Sendo o primeiro exemplo de uma solução de uma equação diferencial estocástica tendo um tempo local como drift ([HARRISON; SHEPP, 1981b](#)).

Considere a equação diferencial estocástica, para um processo $X \in \mathbb{R}$, da forma

$$X_t = x + \sigma B_t + \theta l_t^x, \quad (2.1)$$

onde $B = \{B_t : 0 \leq t \leq T\}$ é o movimento Browniano padrão definido em um espaço de probabilidade, a condição inicial $x \geq 0$ (para o caso $x > 0$ é simétrico), σ é o parâmetro de volatilidade, $\theta \in [-1, 1]$ parâmetro de assimetria, e $l^x = \{l_t^x : 0 \leq t \leq T\}$ é o tempo local em zero.

Dado $T > 0$ fixo, para qualquer intervalo $I \subset [0, T]$, defina

$$l(I, A) := \lambda(\{s \in I; X_s \in A\}), \quad A \subset \mathbb{R},$$

onde λ é a medida de Lebesgue sobre \mathbb{R}^+ , e $l(I, \cdot)$ denota a medida de ocupação em \mathbb{R} de X durante o intervalo de tempo I . Se $l(I, \cdot)$ é absolutamente contínuo em relação à medida de Lebesgue λ em \mathbb{R} , denotamos esta derivada de Radon-Nikodym por

$$l(I, x) = \frac{dl(I, \cdot)}{d\lambda},$$

e chamamos $l(I, x)$ a densidade de ocupação de X no nível x durante o intervalo de tempo I . Escrevemos $l_t(A) = l([0, t], A)$ e $l_t(x) = l([0, t], x)$ para a medida de ocupação e para a densidade (tempo local), respectivamente.

Uma alternativa equivalente para a definição do tempo local $l^x = \{l_t^x : 0 \leq t \leq +\infty\}$ no nível zero de solução (desconhecida) X da equação 2.1, partindo de x , é dado pelo seguinte limite,

$$l_t^x = \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \int_0^t 1_{(-\epsilon, \epsilon)}(X_s) ds.$$

Lembre-se que a solução forte de X da equação diferencial estocástica dada por 2.1 em um determinado espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, em relação ao movimento browniano fixo B e condição inicial independente x sobre este espaço de probabilidade, é um processo estocástico $(X_t, t \geq 0)$ satisfazendo:

1. X é adaptado para a filtragem (\mathcal{F}_t) , onde $\mathcal{F}_t := \sigma(B_s, 0 \leq s \leq t)$;
2. X é um processo contínuo;
3. $\mathbb{P}(X_0 = x) = 1$;
4. com probabilidade um, temos $X_t = x + \sigma B_t + \theta l_t^x, \forall t \geq 0$.

Então, o movimento browniano assimétrico (sBm), $X = \{X_t: 0 \leq t \leq +\infty\}$, pode ser definido como a solução forte (única solução) da equação diferencial estocástica 2.1, (HARRISON; SHEPP, 1981a).

Na literatura, o parâmetro de assimetria é por vezes definido como $\alpha = (\theta + 1)/2$. Esta segunda parametrização é mais conveniente para uma construção alternativa do sBm. Em particular para o desenvolvimento deste documento, a primeira parametrização é usada.

As transições entre os estados são dirigidos pelo seguinte kernel

$$q(t, x, y | \sigma^2, \theta) = \begin{cases} \phi_{t\sigma^2}(y - x) + \theta \phi_{t\sigma^2}(y + x) & \text{si } y > 0 \\ \phi_{t\sigma^2}(y - x)(1 - \theta) & \text{c.o.c.} \end{cases},$$

quando $x > 0$,

$$q(t, x, y | \sigma^2, \theta) = \begin{cases} \phi_{t\sigma^2}(y - x) - \theta \phi_{t\sigma^2}(y + x) & \text{si } y \leq 0 \\ \phi_{t\sigma^2}(y - x)(1 + \theta) & \text{c.o.c.} \end{cases},$$

quando $x < 0$, e quando $x = 0$

$$q(t, 0, y | \sigma^2, \theta) = \begin{cases} \frac{1}{2} \phi_{t\sigma^2}(y)(1 - \theta) & \text{si } y \leq 0 \\ \frac{1}{2} \phi_{t\sigma^2}(y)(1 + \theta) & \text{c.o.c.} \end{cases},$$

onde $\phi_{t\sigma^2}$ denota a densidade de uma variável aleatória Gaussiana centrada com variância $t\sigma^2$. Ver (LEJAY; MORDECKI; TORRES, 2014) por exemplo.

Em particular, se o parâmetro θ é positivo, existe uma assimetria à direita, se θ é zero uma distribuição normal é recuperada com média zero e variância t , e finalmente, se θ é negativo, existe uma assimetria à esquerda.

Uma alternativa para escrever a densidade de transição do sBm, é

$$p(t, x, y) = \begin{cases} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}} + \frac{\theta}{\sqrt{2\pi t}} e^{-\frac{(x+y)^2}{2t}} & \text{if } y > 0 \text{ and } x \geq 0 \\ \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}} - \frac{\theta}{\sqrt{2\pi t}} e^{-\frac{(x+y)^2}{2t}} & \text{if } y < 0 \text{ and } x \leq 0 \\ \frac{1}{\sqrt{2\pi t}} (1 + \theta) e^{-\frac{(x-y)^2}{2t}} & \text{if } y > 0 \text{ and } x \leq 0 \\ \frac{1}{\sqrt{2\pi t}} (1 - \theta) e^{-\frac{(x-y)^2}{2t}} & \text{if } y < 0 \text{ and } x \geq 0. \end{cases}$$

2.1.1 Densidade do sBm

A densidade do processo sBm, em cada instante t , é dado por:

$$f(x, t, \theta) = \frac{1}{\sqrt{2\pi t}} e^{\frac{-x^2}{2t}} [1 + \theta \text{signo}(x)]. \quad (2.2)$$

Propriedades do sBm $\{X_t\}_{t \geq 0}$

1. Esperança:

$$\mathbb{E}(X_t) = \frac{\sqrt{2t}}{\sqrt{\pi}} \theta.$$

2. A variância:

$$\text{Var}(X_t) = t \left(1 - \frac{2\theta^2}{\pi} \right).$$

3. A função Característica do sBm:

$$\varphi_{X_t}(k) = e^{-\frac{tk^2}{2}} \left[1 + \frac{\sqrt{2i}\theta k}{\sqrt{\pi t}} {}_1F_1 \left(\frac{1}{2}; \frac{3}{2}; \frac{k^2 t}{2} \right) \right],$$

onde ${}_1F_1$ é a série hipergeométrica generalizada ([BUCHHOLZ, 2013](#)).

2.2 Processo de Lévy

Um processo estocástico $\{X_t\}_{t \geq 0}$ se diz *processo de Lévy* sobre um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ se verifica as seguintes condições:

1. $\{B_t\}_{t \geq 0}$ tem incrementos independentes;
2. $\mathbb{P}(X_0 = 0) = 1$
3. O processo de Lévy $\{X_t\}_{t \geq 0}$ tem incrementos independentes.
4. O processo de Lévy $\{X_t\}_{t \geq 0}$ tem *incrementos estacionários*.
5. Para tudo $\epsilon > 0$ e $s \geq 0$,

$$\lim_{t \rightarrow s} \mathbb{P}(|X_t - X_s| > \epsilon) = 0,$$

isto é, o processo $\{X_t\}_{t \geq 0}$ é *estocasticamente contínuo*.

Para um tratamento completo dos processos de Lévy veja ([APPLEBAUM, 2009](#)).

2.2.1 Processos Subordinados

A ideia de processo subordinado é introduzida por (BOCHNER, 1949). Para um desenvolvimento completo de suas propriedades, veja por exemplo (SATO, 1999).

Definição 2. Um processo Lévy unidimensional não decrescente (a.s.) é chamado de subordinado.

Tal processo pode ser interpretado como um modelo aleatório da evolução do tempo, já que se $T = \{T(t), t \geq 0\}$ é um processo subordinado, se tem

$$T(t) \geq 0 \quad \text{a.s. para cada } t > 0, \quad (2.3)$$

y

$$T(t_1) \leq T(t_2) \quad \text{a.s. quando } t_1 \leq t_2. \quad (2.4)$$

2.2.1.0.1 Subordinados.

Neste trabalho os processos descritos a seguir serão utilizados como subordinados.

1. **Processo Gama:** O Processo $\{\gamma_t\}_{t \geq 0}$ é chamado de processos Gama, se a sua densidade para cada $t \geq 0$, é dada por

$$f_{\gamma_t}(x) = \frac{b^{at}}{\Gamma(at)} x^{at-1} e^{-bx}, \quad \text{con } \alpha, \beta > 0, \quad (2.5)$$

onde $\Gamma(t)$ é a *função Gama* em relação a t .

Algumas propriedades do processo γ_t são:

- a) A esperança do processo γ_t ,

$$\mathbb{E}(\gamma_t) = \frac{at}{b}, \quad t > 0.$$

- b) A variância do processo γ_t ,

$$Var(\gamma_t) = \frac{at}{b^2}.$$

- c) A função característica, em cada $t > 0$, para o processo γ_t é

$$\varphi_{\gamma_t}(k) = \left(1 - \frac{ik}{b}\right)^{-at}.$$

- d) A covariância do processo Gama,

$$Cov(\gamma_s, \gamma_t) = \frac{a}{b^2} \min\{s, t\}.$$

e) A correlação entre γ_s e γ_t

$$Corr(\gamma_s, \gamma_t) = \sqrt{\frac{s}{t}},$$

para $s < t$.

Um aspecto relevante para a simulação do processo γ_t é que tem incrementos independentes, isto é,

$$\gamma_{t+h} - \gamma_t \sim f_{\gamma_h}(x) = \frac{b^{ah}}{\Gamma(ah)} x^{ah-1} e^{-bx}, \quad h > 0.$$

Além disso, se considerarmos $a = b = 1/2$ os incrementos têm distribuição qui-quadrado com h graus de liberdade (χ_h^2), isso é,

$$\gamma_{t+h} - \gamma_t \sim f_{\gamma_h}(x) = \frac{1}{2^{\frac{h}{2}} \Gamma\left(\frac{h}{2}\right)} x^{\frac{h}{2}-1} e^{-\frac{x}{2}}.$$

2. Processo Gaussiano Inverso: O processo $\{\tau_t\}_{t \geq 0}$ é chamado Gaussiano inverso ($\tau_t \sim IG(\alpha, \gamma t)$, com $\alpha, \gamma > 0$), se para cada τ_t com $t \geq 0$ sua densidade é dada por

$$f_{\tau_t}(x) = \frac{\gamma t e^{\gamma t \alpha}}{\sqrt{2\pi x^3}} e^{-\frac{1}{2}\left(\frac{t^2 \gamma^2}{x} + \alpha^2 x\right)}, \quad x \geq 0, \quad (2.6)$$

Algumas propriedades do processo τ_t são:

a) A esperança do processo τ_t ,

$$\mathbb{E}(\tau_t) = \frac{\gamma t}{\alpha}, \quad t > 0.$$

b) A variância do processo τ_t ,

$$Var(\tau_t) = \frac{\gamma t}{\alpha^3}.$$

c) A função característica, em cada $t > 0$, para o processo τ_t é

$$\varphi_{\tau_t}(k) = \sqrt{\frac{2t\gamma}{\pi}} e^{\gamma t \alpha} (\alpha^2 - 2ik)^{1/4} K_{1/4}(t\gamma\sqrt{\alpha^2 - 2ik}).$$

d) A covariância do processo Gaussiano inverso,

$$Cov(\tau_s, \tau_t) = \frac{\gamma}{\alpha^3} \min\{s, t\}.$$

e) A correlação entre τ_s e τ_t

$$Corr(\tau_s, \tau_t) = \sqrt{\frac{s}{t}},$$

para $s < t$.

Um aspecto importante do processo τ_t é que tem aumentos independentes, isto é,

$$\tau_{t+h} - \tau_t \sim f_{\tau_h}(x) = \frac{\gamma h e^{\gamma h \alpha}}{\sqrt{2\pi x^3}} e^{-\frac{1}{2}\left(\frac{h^2 \gamma^2}{x} + \alpha^2 x\right)}, \quad h > 0.$$

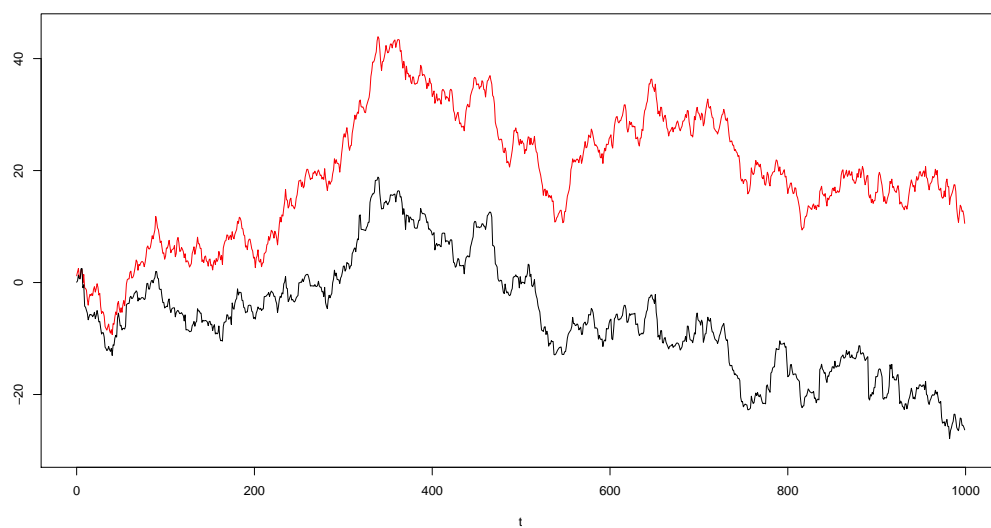


Figura 1 – Em vermelho o movimento Browniano padrão, e em preto o movimento Browniano com processo subordinado T .

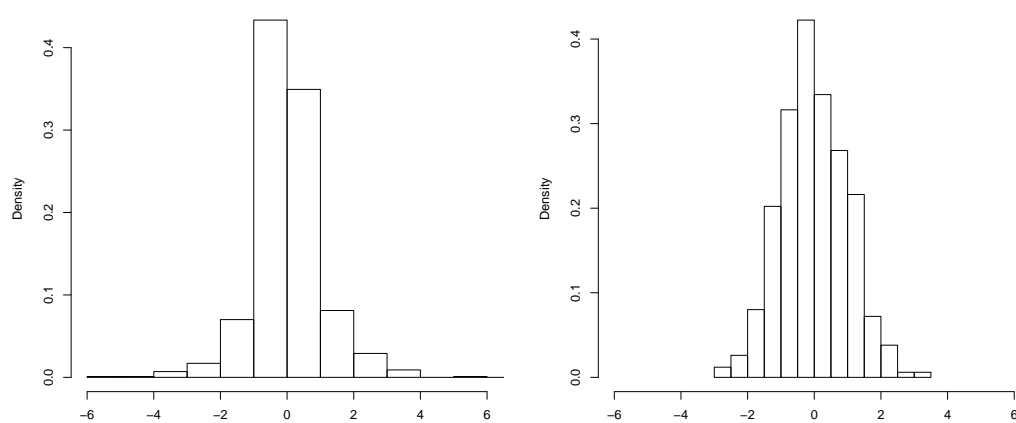


Figura 2 – À esquerda, incrementos do movimento Browniano com processo subordinado T , e à direita, os incrementos do movimento Browniano padrão.

2.3 Modelo de Black-Scholes

A partir de os trabalhos de (BACHELIER, 1900), são usados os processos em tempo contínuo para modelar o comportamento do preço de um ativo. Black, F. e Scholes M. (BLACK; SCHOLES, 1973) desenvolveram a fórmula que é a base de todo o mercado de opções. Uma opção é entendida como o pagamento para possuir o direito, mas não a obrigação com data de validade fixa, em que o contratante, tem a opção de comprar ou vender ativos em valor que é chamado o preço de exercício.

O modelo é uma das maiores conquistas em matemática financeira, descreve o valor de uma opção ao longo do tempo e torna possível encontrar explicitamente o valor de uma opção sob certas condições ou suposições.

O modelo de Black-Scholes parte de hipóteses semelhantes ao modelo Binomial (COX; ROSS; RUBINSTEIN, 1979) sobre o funcionamento do mercado e acrescenta algumas hipóteses específicas sobre a evolução do preço subjacente.

2.3.1 Modelo

Consideramos agora o movimento Browniano geométrico, um modelo de matemática financeira também conhecido como modelo de Black-Scholes. Onde o preço S_t ao tempo t , de um ativo é:

$$S_t = S_0 \exp[\mu t + \sigma W(t)], \quad (2.7)$$

onde $\mu, \sigma^2 > 0$ são constantes fixas e $W = \{W(t), t \geq 0\}$ é o movimento Browniano. Note também que 2.7 é a solução de Itô da equação diferencial estocástica,

$$dS(t) = S(t)((\mu + \frac{1}{2}\sigma^2)dt + \sigma dW(t)).$$

O retorno do processo é

$$X_t = \log \left(\frac{S_t}{S_{t-1}} \right) = \mu + \sigma(W(t) - W(t-1)). \quad (2.8)$$

As seguintes propriedades são conseqüências diretas das expressões 2.7, 2.8.

- i.- O processo $\{X_t\}_{t \geq 0}$ tem distribuição estacionária Gaussiana.
- ii.- $\{X_t\}$ tem incrementos independentes.
- iii.- $Cov((X_{t+k} - \mu)^2, (X_t - \mu)^2) = 0$, isto é, os retornos ao quadrado são independentes.
- iv.- $Cov(|X_{t+k} - \mu|, |X_t - \mu|) = 0$, isto é, os retornos absolutos não são correlacionados.

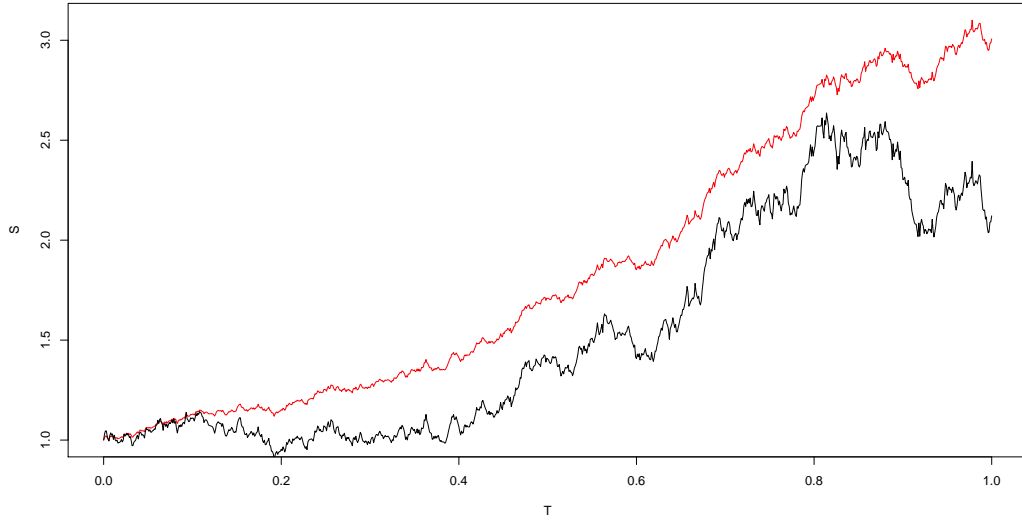


Figura 3 – O movimento Browniano geométrico, considerando em preto $\mu = 1$, $\sigma = 0.5$ e em vermelho $\mu = 0.5$, $\sigma = 0.5$

Se denotarmos a $P(t)$ $t \geq 0$, como o preço de um ativo ao tempo t , $X(t)$ são chamados de retornos, ao tempo t e escala τ e representa a variação relativa do preço de $t - \tau$ até t ,

$$X_\tau(t) = \frac{P(t) - P(t - \tau)}{\tau},$$

se τ é suficientemente pequeno, temos aproximadamente

$$X_\tau(t) \simeq \log P(t) - \log P(t - \tau).$$

2.3.2 Distribuição tipo G

A família de distribuições de probabilidade do tipo G, é gerada a partir do produto entre uma família de distribuições normais e outra de distribuições infinitamente divisíveis (i.d.) definidas em reais positivos, e também conhecidas como distribuições normais com mistura de variância e média (ROSINSKI, 1991). Esta família de distribuições é discutida por (MARCUS, 1987), estudada por (ROSINSKI, 1991).

Definição 3. *Seja X uma variável aleatória contínua do tipo G definida como:*

$$X \stackrel{d}{=} \mu + \beta T^2 + TW, \quad (2.9)$$

onde T e W são variáveis aleatórias independentes, W tem distribuição normal com média zero e variância constante, T é não-negativo e pertence à classe de distribuições infinitamente divisíveis, μ é o parâmetro de localização, o parâmetro β refere-se à assimetria da variável resultante, com $\beta = 0$ a variável aleatória X é simétrica, $\beta < 0$ ou $\beta > 0$ a variável aleatória X é assimétrica negativa ou positiva respectivamente.

Se T tem uma distribuição Gaussiana Inversa Generalizada (GIG) a variável X tem distribuição Hiperbólica Generalizada (BARNDORFF-NIELSEN, 1977). Esta distribuição particular tem três parâmetros $GIG(\lambda, \gamma, \epsilon)$ foi usado por (SICHEL, 1974) ao construir misturas de uma distribuição de Poisson e foi popularizado por (BARNDORFF-NIELSEN, 1977), que obteve uma distribuição hiperbólica generalizada ao fazer misturas de uma distribuição normal com média zero e variância constante, a distribuição inversa Gaussiana é obtida quando se considera $\lambda = -1/2$. Várias propriedades probabilísticas da distribuição GIG como divisibilidade infinita e auto-decomposição são demonstradas por (BARNDORFF-NIELSEN; HALGREEN, 1977) e (HALGREEN, 1979). A distribuição GIG é semelhante à distribuição von Mises (BARNDORFF-NIELSEN, 1977) e compartilha muitas propriedades dessa distribuição. Tais propriedades especiais da distribuição GIG não são tratados neste trabalho, ver (FOLKS; CHHIKARA, 1978) para uma revisão mais detalhada. O conceito de infinita divisibilidade de uma variável aleatória foi introduzido em 1929 por Bruno de Finetti que deu o contexto matemático de processos estocásticos, com incrementos estacionários e independentes (sii) seu desenvolvimento foi concluído com a forma mais geral da representação canônica das funções características de variáveis aleatórias infinitamente divisíveis ((STEUTEL; HARN, 2003)).

Este trabalho lida com diferentes estruturas para T , considerado como um processo de Lévy.

2.4 FATGBM

O Movimento Browniano Geométrico com Tempo Fractal (FATGBM por suas siglas em inglês) proposto originalmente por (HEYDE, 1999), descreve um modelo semelhante ao modelo de Black-Sholes para o preço de opções, usando movimento Browniano geométrico com tempo fractal ou subordinado aleatório, para duas construções alternativas deste processo ver (KERSS; LEONENKO; SIKORSKII, 2014). No modelo FATGBM o preço P_t ao tempo t , de um ativo é:

$$P_t = P_0 e^{\mu t + \theta T_t + \sigma W(T_t)}, \quad t \geq 0, P_0 > 0, \quad (2.10)$$

onde $\mu \in \mathbb{R}$, $\theta \in \mathbb{R}$ y $\sigma > 0$ são os parâmetros de drift, assimetria e volatilidade, respectivamente, $W(t), t \geq 0$ é o movimento browniano, independente da "atividade de tempo", $\{T_t\}$. $\{T_t, t = 0, 1, 2, \dots, T_0 = 0\}$ é um processo estritamente positivo, não decrescente, com incrementos estacionários, mas não necessariamente independentes. Os incrementos sob a unidade de tempo são dados por

$$\tau = T_t - T_{t-1}, \quad t = 1, 2, \dots$$

(HEYDE; LEONENKO, 2005) No contexto financeiro, esta "atividade de tempo" $\{T_t\}$ pode ser interpretada como o tempo durante o qual os preços de mercado evoluem (HOWISON;

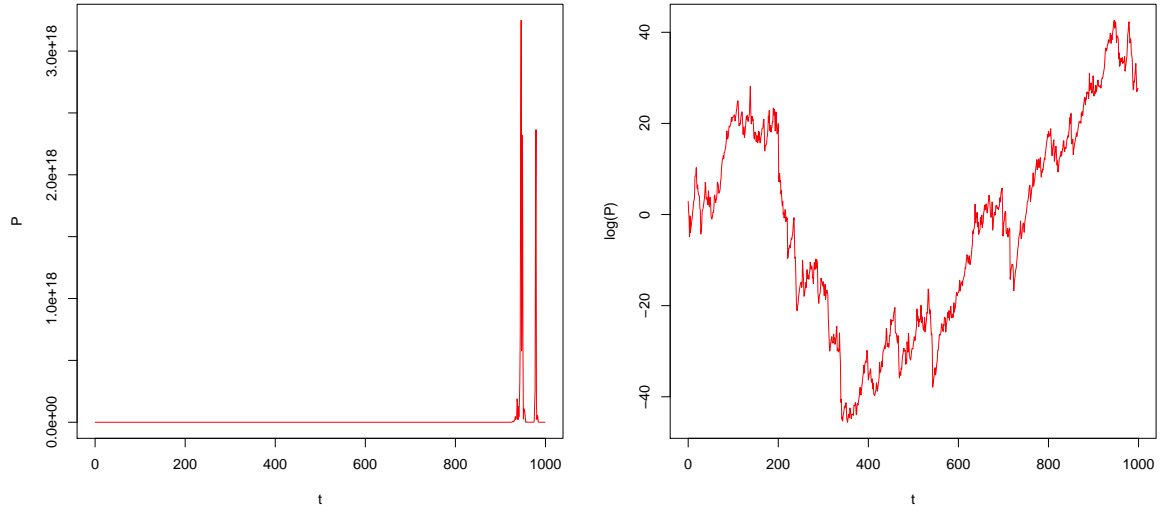


Figura 4 – À esquerda, simulação de FATGBM com $\mu = 1$, $\theta = -1$ e $\sigma = 1$, e à direita, o \log de FATGBM.

(LAMPER, 2001). Quanto mais informação é lançada no mercado, mais rápido flui o $\{T_t\}$. Note que se $T_t = t$, então a equação descreve o modelo clássico de Black-Scholes. Os retornos correspondentes são dados por,

$$X_\tau = \mu + \theta\tau_t + \sigma\tau_t^{1/2}W(1), \quad t = 1, 2, \dots \quad (2.11)$$

O papel que desempenha $\{T_t\}$ no modelo é crucial, pois determina a distribuição de $\{X_t\}$ e sua estrutura de correlação.

2.4.1 Distribuição Teórica do Modelo.

A partir da expressão 2.11 podemos ver que a distribuição condicional de X_t dado $\tau_t = h$ é normal, com média $\mu + \theta h$, e variância $\sigma^2 h$, isso é,

$$X_t | \tau_t = h \sim N(\mu + \theta h, \sigma^2 h). \quad (2.12)$$

Depois do qual é possível deduzir algumas propriedades do processo.

1. A esperança do processo:

$$\mathbb{E}(X_t) = \mu + \theta\mathbb{E}(h).$$

2. A variação para o processo em cada tempo $\tau_t = h$

$$\text{Var}(X_t) = \sigma^2\mathbb{E}(h) + \theta^2\text{Var}(h).$$

3. O coeficiente de assimetria da distribuição de X_t é,

$$\gamma_1 = \frac{\theta((3\mu\theta + 3\sigma^2)Var(h) + \theta^2\mathbb{E}(h - \mathbb{E}(h))^3)}{(\sigma^2\mathbb{E}(h) + \theta^2Var(h))^{\frac{3}{2}}}.$$

4. O coeficiente de curtose de X_t ,

$$\begin{aligned} \gamma_2 = & [6(2\mu\theta^2 + \theta\sigma^2 - 2\mu\theta^3)\mathbb{E}^3(h) + 3(4\mu\theta\sigma^2 - 4\theta\mu^3 + \sigma^4)\mathbb{E}(h^2) + 6\theta^2\sigma^2\mathbb{E}(h^3) \\ & - 12\theta^2\mu^2\mathbb{E}(h)\mathbb{E}(h^2) + \theta^4\mathbb{E}(h - \mathbb{E}(h))^4](\sigma^2\mathbb{E}(h) + \theta^2Var(h))^{-2}. \end{aligned}$$

Os coeficientes de assimetria e curtose existem se o terceiro e quarto momentos de $\{\tau_t\}$ são finitos respectivamente. Além disso, observe que se $\theta = 0$ o modelo é simétrico.

2.4.2 Distribuição dos Incrementos do Processo Subordinados

A seguir se mostram algumas distribuições possíveis para τ_t . Em cada caso, se mostra a distribuição de X_t .

2.4.2.1 Subordinado com distribuição Gama Inversa

Se τ_t tem uma distribuição Gama inversa, com parâmetros $\alpha, \beta > 0$ e função de densidade dada por,

$$f_{R\Gamma}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\beta/x}, x > 0. \quad (2.13)$$

Então, a densidade marginal de X_t para $\theta = 0$ é dada por,

$$f(x) = \frac{\beta^\alpha \Gamma(\alpha + 1/2)}{\sqrt{2\pi}\sigma \Gamma(\alpha) \left(\frac{(x-\mu)^2 + 2\sigma^2\beta}{2\sigma^2} \right)^{\alpha+1/2}}, \quad x \in \mathbb{R}.$$

Por outro lado, se $\theta \neq 0$ então a densidade de X_t é

$$\begin{aligned} f(x) = & \frac{\beta^\alpha e^{\frac{\theta(x-\mu)}{\sigma^2}}}{\sqrt{2\pi}\sigma^2 \Gamma(\alpha)} 2 \left(\frac{(x-\mu)^2 + 2\beta\sigma^2}{\theta^2} \right)^{-\frac{\alpha+1/2}{2}} \times \\ & \times K_{-\alpha-1/2} \left(\frac{\sqrt{\theta^2((x-\mu)^2 + 2\beta\sigma^2)}}{\sigma^2} \right), x \in \mathbb{R}. \end{aligned}$$

Onde K_ν é a função Bessel modificada de terceira ordem, ou a função McDonalds com índice ν . Dada por,

$$K_\nu = \int_{\mathbb{R}^+} e^{-x \cosh(t)} \cosh(\nu t) dt.$$

2.4.2.2 Subordinado com distribuição de Gama

Se τ_t é distribuído Gama com parâmetros $\alpha, \beta > 0$, isso é,

$$f_{\Gamma}(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, x > 0. \quad (2.14)$$

Então a densidade marginal de X_t , quando $\theta = 0$, é dada por,

$$f(x) = \sqrt{\frac{2}{\pi}} \frac{\beta^{\alpha}}{\sigma \Gamma(\alpha)} \left(\frac{|x - \mu|}{\sigma \sqrt{2\beta}} \right)^{\alpha-1/2} K_{\alpha-1/2} \left(\frac{|x - \mu| \sqrt{2\beta}}{\sigma} \right), \quad x \in \mathbb{R}.$$

E se $\theta \neq 0$, então

$$\begin{aligned} f(x) = & \sqrt{\frac{2}{\pi}} \frac{\beta^{\alpha} e^{\frac{\theta(x-\mu)}{\sigma^2}}}{\sigma \Gamma(\alpha)} \left(\frac{|x - \mu|}{\sqrt{\theta^2 + 2\beta\sigma^2}} \right)^{\alpha-1/2} \times \\ & \times K_{\alpha-1/2} \left(\frac{|x - \mu| \sqrt{\theta^2 + 2\beta\sigma^2}}{\sigma^2} \right), \quad x \in \mathbb{R} \end{aligned} \quad (2.15)$$

2.4.2.3 Subordinado com distribuição Gaussiana Inversa

Se τ_t tem distribuição $IG(\alpha, \gamma)$ com parâmetros $\alpha > 0$, $\gamma \geq 0$, então a sua função de densidade é dada por,

$$f_{IG}(x) = \frac{\gamma e^{\gamma\alpha}}{\sqrt{2\pi} x^3} e^{-\frac{1}{2} \left(\frac{\gamma^2}{x} + \alpha^2 x \right)}, x > 0. \quad (2.16)$$

Quando $\theta = 0$ a densidade marginal de X_t é dada por,

$$f(x) = \frac{\gamma}{\pi} e^{\gamma\alpha} \frac{\alpha}{\sqrt{(x - \mu)^2 + \sigma^2 \gamma^2}} K_1 \left(\frac{\alpha \sqrt{(x - \mu)^2 + \sigma^2 \gamma^2}}{\sigma} \right), \quad x \in \mathbb{R}.$$

Por outro lado, se $\theta \neq 0$, então

$$\begin{aligned} f(x) = & \frac{\gamma}{\pi \sigma} e^{\gamma\alpha + \frac{\theta(x-\mu)}{\sigma^2}} \sqrt{\frac{\theta^2 + \sigma^2 \alpha^2}{(x - \mu)^2 + \sigma^2 \gamma^2}} \times \\ & \times K_1 \left(\frac{\sqrt{(\theta^2 + \sigma^2 \alpha^2)((x - \mu)^2 + \sigma^2 \gamma^2)}}{\sigma^2} \right), x \in \mathbb{R}, \end{aligned}$$

onde K_1 é a função de Bessel modificada do terceiro tipo com índice 1.

2.4.3 Estrutura de dependência do Modelo

Propriedades da estrutura de covariância do processo $\{X_t\}$ em termos do processo subordinado, assumindo momentos finitos de segundo ordem como condição necessária. Para o inteiro $k > 1$, temos,

$$\begin{aligned} Cov(X_t, X_{t+k}) &= Cov(\theta \tau_t + \sigma \tau_t^{\frac{1}{2}} W_1(1), \theta \tau_{t+k} + \sigma \tau_{t+k}^{\frac{1}{2}} W_2(1)), \\ &= \theta^2 (E \tau_t \tau_{t+k} - E \tau_t E \tau_{t+k}), \\ &= \theta^2 Cov(\tau_t, \tau_{t+k}), \end{aligned}$$

onde W_1 y W_2 são movimentos brownianos independentes. Note também que se $\theta = 0$, então $Cov(X_t, X_{t+k}) = 0$.

Também

$$\begin{aligned} Cov(X_t^2, X_{t+k}^2) &= Cov((\mu + \theta\tau_t + \sigma\tau_t^{\frac{1}{2}}W_1(1))^2, (\mu + \theta\tau_{t+k} + \sigma\tau_{t+k}^{\frac{1}{2}}W_2(1))^2), \\ &= (\sigma^4 + 4\theta^2\mu^2 + 4\theta\mu\sigma^2)Cov(\tau_t, \tau_{t+k}) + \theta^4Cov(\tau_t^2, \tau_{t+k}^2) \\ &\quad + (\theta^2\sigma^2 + 2\theta^3\mu)(Cov(\tau_t^2, \tau_{t+k}) + Cov(\tau_t, \tau_{t+k}^2)), \end{aligned}$$

que é reduzido a

$$Cov(X_t^2, X_{t+k}^2) = \sigma^4Cov(\tau_t, \tau_{t+k}),$$

quando $\theta = 0$.

Parte III

Inferência Bayesiana

3 Elementos de Inferência Bayesiana

3.1 Inferência Bayesiana

Na inferência clássica, os parâmetros de um modelo são vistos como quantidades fixas, mesmo que sejam desconhecidos. Em contraste, na inferência bayesiana, o conhecimento que temos sobre os parâmetros do modelo pode ser quantificado através de uma distribuição de probabilidade. Essa distribuição de probabilidade é chamada de distribuição a priori do parâmetro, e ela contém toda a informação que o pesquisador tem sobre os parâmetros. As inferências para os parâmetros do modelo são obtidas através da distribuição a posteriori de cada parâmetro.

Seja θ um vetor de parâmetros a ser estimado, com espaço paramétrico $\Theta \subseteq \mathbb{R}^n$. Então, pelo teorema de Bayes, temos a seguinte distribuição de probabilidade a posteriori para θ ,

$$\pi(\theta|y) = \frac{\pi(\theta)f(y|\theta)}{\int_{\Theta} \pi(\theta)f(y|\theta)d\theta}. \quad (3.1)$$

onde θ é contínuo, $\pi(\theta)$ é a distribuição a priori conjunta para θ e $f(y|\theta) = L(\theta)$ é a função de verossimilhança de θ . A forma usual do teorema de Bayes é

$$\pi(\theta|y) \propto \pi(\theta)f(y|\theta). \quad (3.2)$$

Note que omitindo $f(y)$, a igualdade em 3.1 é substituída pela proporcionalidade. É uma forma simplificada do teorema de Bayes.

3.1.1 Teoria de Decisão

Constituem os elementos fundamentais da teoria da decisão:

1. Um modelo estatístico \mathbb{P} sobre $(\mathcal{X}, \mathcal{F}_{\mathcal{X}})$. Em particular, considerando um modelo paramétrico,

$$\mathcal{P} = \{P_{\theta} : \theta \in \Theta\},$$

como a família de distribuições de um vetor aleatório X , que assume valores em $(\mathcal{X}, \mathcal{T})$.

2. Um espaço mensurável $(\mathbb{A}, \mathcal{F}_{\mathbb{A}})$, chamado espaço de ações. O conjunto \mathbb{A} é o conjunto de decisões estatísticas. (Conclusões sobre θ ou P_{θ} tomadas com base em as observações)

3. Uma função de perda $L : \Theta \times \mathbb{A} \rightarrow \mathbb{R}^+$ tal que $L(\theta, \cdot)$ é Borel mensurável para tudo $\theta \in \Theta$. O número $L(\theta, a)$ representa a perda em escolher a ação a para cada θ .

Definidos esses três elementos, apresentamos as ferramentas utilizadas na tomada de decisão.

Definição 4. Uma regra de decisão é uma função $\delta : \mathcal{X} \rightarrow \mathbb{A}$ tal que $\delta(x) = a$, isto é, todo valor possível de $x \in \mathcal{X}$, indica que ação a deve ser tomada se X assume o valor x .

Definição 5. A função de perda associada a uma regra de decisão δ é dado por,

$$L(\theta, \delta(x)) = \int_{\mathbb{A}} L(\theta, a) d\delta(x)(a).$$

Perdas comumente usadas são:

$$\begin{aligned} L(\theta, a) &= (g(\theta) - a)^2; \\ L(\theta, a) &= |g(\theta) - a|; \\ L(\theta, a) &= k(\theta)|g(\theta) - a|^p, \quad k(\theta) \geq 0. \end{aligned}$$

Definição 6. O risco a posteriori de uma regra de decisão δ é definido por:

$$\begin{aligned} r(\delta|x) &= \mathbb{E}(L(\theta, \delta(x))|X = x) \\ &= \int_{\Omega} L(\theta, \delta(x)) d\mathbb{P}_{\theta|x}(\theta|x). \end{aligned}$$

A comparação de duas regras de decisão na perspectiva bayesiana é baseada em na escolha de uma com o menor risco após a observação dos dados.

Definição 7. Uma regra de decisão δ_0 é chamada de regra Bayesian se,

$$\begin{aligned} r(\delta_0|x) &< \infty, \quad \forall x \in \mathcal{X} \\ r(\delta_0|x) &< r(\delta|x), \quad \forall x \in \mathcal{X}, \end{aligned}$$

e qualquer outra regra de decisão δ .

De acordo com o critério de Bayes, $\delta(\theta)$ é a regra de Bayes formal com relação a θ se o risco a posteriori é o mínimo possível, isto é, se

$$r(\delta|x) = \mathbb{E}(L(\theta, \delta(x))|X = x) = \min_{\hat{\theta} \in \Theta} \mathbb{E}(L(\theta, \hat{\theta})|X = x), \quad \forall x \in \mathcal{X}.$$

Para mais detalhes, veja por exemplo (DEGROOT; SCHERVISH, 2012; SCHERVISH, 2012).

3.1.2 Estimativa por Regiões

Um resumo de $\pi(\theta|x)$ mais informativos do que qualquer estimativa pontual é obtido a partir de uma região de Θ que contenha uma parte substancial da massa de probabilidade a posteriori.

Definição 8. $R(x)$ é uma região de credibilidade γ para θ se

$$\mathbb{P}[\theta \in R(x)|x] \equiv \int_{R(x)} d\mathbb{P}_{\theta|x}(\theta|x) = \int_{R(x)} \pi(\theta|x) d\theta \geq \gamma.$$

Qualquer região de credibilidade é definida numericamente, isso é, não aleatório, e admite uma interpretação probabilística direta e inequívoca em contraste com uma região de confiança clássica.

Dadas as inúmeras regiões de credibilidade com o mesmo grau de credibilidade γ , obviamente, estamos interessados em selecionar aqueles que incluem todos os valores de θ “credíveis” a posteriori, satisfazendo a seguinte relação,

$$\pi(\theta_1|x) \geq \pi(\theta_2|x), \quad \forall \theta_1 \in R(x), \theta_2 \notin R(x),$$

Assim definimos o seguinte,

Definição 9. $R(x)$ é uma região de credibilidade γ com densidade (probabilidade) a posteriori máxima (abreviado, região de HPD) se

$$\sup_{\theta \notin R(x)} \pi(\theta|x) \leq c_\gamma \leq \inf_{\theta \in R(x)} \pi(\theta|x),$$

para algum c_γ tal que $\mathbb{P}[\theta \in R(x)|x] \geq \gamma$, ou equivalentemente, se

$$R(x) = \{\theta : \pi(\theta|x) \geq c_\gamma\},$$

com $c_\gamma > 0$ a maior constante tal que

$$\int_{R(x)} \pi(\theta|x) \lambda(d\theta) \geq \gamma. \quad (3.3)$$

As regiões de credibilidade HPD não são invariantes para transformações paramétricas não lineares, isso é, se $R(x)$ é HPD para θ , a região transformada para a função biunívoca $\alpha = \varphi(\theta)$ tem a mesma credibilidade, mas não é HPD se $\varphi(\cdot)$

As integrais a ser resolvidas em 3.1 e 3.3 podem ser difíceis de resolver. Neste caso, alguns métodos numéricos de aproximação de integrais devem ser utilizados. Em particular, o método que usaremos para resolver este tipo de integrais é o método de Monte Carlo em cadeias de Markov (MCMC). Que também é usado como um método de simulação para obter amostras da distribuição a posteriori de interesse, veja por exemplo (VOSS, 2013).

3.1.3 Teste de Hipótese Bayesiana

Dentro das diversas alternativas de procedimentos de testes para hipóteses precisas, utilizaremos aquela construída por Pereira and Stern (BRAGANÇA-PEREIRA; STERN, 1999) e (PEREIRA et al., 2008), denominada *Full Bayesian Significance Test* (FBST), comparando-a com outros procedimentos de seleção de modelos usualmente adotados.

Esta metodologia tem sido aplicada com sucesso em contextos onde a metodologia clássica não é a mais apropriada ou é simplesmente inexistente, como discutido nas referências citadas.

Especificamente, no caso de processos de difusão com saltos, (RIFO; TORRES, 2009) estuda o caso de identificação de saltos com aplicações a dados meteorológicos, e no caso de processos de difusão assimétricos, (TORRES; TUDOR, 2009) aplica a metodologia a dados financeiros e (BARAHONA et al., 2016) a dados de localização de leões marinhos da América do Sul.

Outros trabalhos têm aplicado este procedimento em diversos contextos: (BERNARDINI; RIFO, 2011) estuda o parâmetro de forma na distribuição de Valores Extremos Generalizada, (RODRIGUES, 2003) analisa o processo de Poisson inflacionado em zero, (RIFO; GONZÁLEZ-LÓPEZ, 2012) modela dependência caudal via cópulas logísticas assimétricas para duas variáveis, (LOSCHI; SANTOS; ARELLANO-VALLE, 2010) estende a metodologia a um critério de decisão ponderado relacionando o FBST e fator de Bayes, entre outros.

Seja Ω o espaço paramétrico definido para o problema em estudo, e denotado por $\pi(\theta)$ a densidade a priori em Ω , por $L_d(\theta)$, a função de verossimilhança para θ depois de observar os dados \mathbf{d} , e por $\pi(\theta|\mathbf{d})$ a densidade a posteriori de θ dada a amostra d , relacionados por

$$\pi(\theta|\mathbf{d}) \propto \pi(\theta) \cdot L_d(\theta).$$

Uma hipótese simples é uma sub-variedade Ω_0 de Ω tal que $\dim(\Omega_0) < \dim(\Omega)$. Isto implica que a probabilidade a posteriori de Ω_0 é igual a 0 no caso em que a distribuição a posteriori seja absolutamente contínua. Uma medida de evidência neste caso é definida pelo teste de significância FBST, definido a partir da probabilidade a posteriori de uma espécie de região crítica do espaço paramétrico.

Mais precisamente, vamos definir o conjunto de tangente a Ω_0 como o conjunto

$$T_0 = \{ \theta \in \Omega : \pi(\theta|\mathbf{d}) > \pi_0, \},$$

onde

$$\pi_0 = \sup_{\Omega_0} \pi(\theta|\mathbf{d}),$$

definido pelas curvas de nível da densidade a posteriori. Em palavras, o conjunto tangente considera todos os pontos mais “prováveis” a posteriori do que aqueles em Ω_0 . A medida de evidência e —valor em favor da hipótese nula é definida como

$$ev(\Omega_0) = 1 - \int_{T_0} \pi(\theta|\mathbf{d}) d\theta. \quad (3.4)$$

Desta forma, se o conjunto tangente tem uma alta probabilidade a posteriori, a evidência em favor de Ω_0 é pequena; se tiver uma baixa probabilidade a posteriori, a evidência contra Ω_0 é pequena.

Tal como é descrito em (BERNARDINI; RIFO, 2011), pela continuidade da distribuição a posteriori em Ω , existe uma representação de $\pi(\omega|\mathbf{d})$ tal que

$$\pi_0 = \limsup_{\epsilon \rightarrow 0} \{ \pi(\omega|\mathbf{d}) : \omega \in V_0(\epsilon) \}, \quad (3.5)$$

onde $V_0(\epsilon) = \{ \omega \in \Omega : \text{dist}(\omega, \Omega_0) < \epsilon \}$ é uma vizinhança de Ω_0 .

Esta propriedade nos permite determinar π_0 sem a necessidade de obter a densidade a priori condicional na hipótese nula, usando o limite 3.5 em seu lugar.

3.1.3.1 Teste de Significância

Sob esta formulação, esta medida nos permite realizar um teste de significância para $H_0: \theta \in \Omega_0$. De fato, consideramos o espaço de decisão $\mathcal{D} = \{a_0, a_1\}$, onde a_1 representa a decisão de rejeitar H_0 , y a_0 , não rejeitar H_0 .

A função de perda $L : \mathcal{D} \times \Omega \rightarrow \mathbb{R}$ é definida por

$$L(a_1, \theta) = \omega_0(1 - \mathbf{1}_{T_0}(\theta)),$$

$$L(a_0, \theta) = \omega_1 + c\mathbf{1}_{T_0}(\theta),$$

com $\omega_0, \omega_1, c > 0$. O valor de c representa o custo adicional de não rejeitar Ω_0 quando de fato $\theta \in T_0$, e portanto, tem densidade a posteriori maior do que qualquer valor em Ω_0 .

De acordo com (MADRUGA; ESTEVES; WECHSLER, 2001), a regra de decisão definida por

$$\text{“rejeitar } \Omega_0 \text{”, se, e somente se } e - \text{valor} < \frac{\omega_1 + c}{\omega_0 + c}, \quad (3.6)$$

minimiza o valor esperado $\mathbb{E}(l(a, \theta)|\mathbf{d})$.

Em outras palavras, a decisão 3.6 entrega um critério em termos do custo de cada decisão incorreta e a probabilidade a posteriori do conjunto tangencial.

Por exemplo, se tomarmos $\omega_1 \ll c$, e reescrevemos as condições em 3.6, obtemos $c/\omega_0 > e/(1-e)$. Ou seja, rejeitamos H_0 se o custo relativo de rejeitar H_1 é maior que a razão de probabilidade contra o conjunto tangencial.

3.1.3.2 Critério de Informação Bayesiana

O Critério de Informação Bayesiana (BIC por suas siglas em inglês) para a hipótese i , é definido como

$$BIC_i = 2l_i + p_i n, \quad (3.7)$$

onde p_i é o número de parâmetros do modelo i e n é o tamanho da amostra. De acordo com este bem conhecido critério, o melhor modelo é o que tem menor BIC , (SPIEGELHALTER et al., 2002).

Nesta tese, usamos as abordagens anteriores para obter medidas de evidência bayesiana para as hipóteses de simetria, sobre o parâmetro θ e σ em sBm.

3.2 Computação Bayesiana Aproximada

A classe de algoritmos computacionais conhecida como Computação Bayesiana Aproximada, ou Approximate Bayesian Computation (ABC), é uma alternativa aos métodos de Monte Carlo baseados em cadeias de Markov, ou MCMC (Markov Chain Monte Carlo). Tais métodos são eficientes no contexto de funções de verossimilhança complexas ou intratáveis analítica e computacionalmente.

Os métodos ABC são usados para aproximar a distribuição a posteriori sem exigir uma expressão explícita para a função de verossimilhança. Estes métodos tornam-se importantes com os trabalhos de (TAVARÉ et al., 1997), (BEAUMONT; ZHANG; BALDING, 2002) e (PRITCHARD et al., 1999) em genética e (CSILLÉRY et al., 2010) em ecologia.

O algoritmo proposto por (PRITCHARD et al., 1999), pode ser resumido da seguinte forma: dada uma amostra D de um espaço \mathcal{D} ,

Algoritmo 1 – Computação Bayesiana Aproximada (ABC)

- A1 Gerar θ desde a distribuição a priori $\pi(\cdot)$
 - A2 Simular D^* desde o modelo $f(\cdot|\theta)$ com parâmetro θ , então calcule $T^* = T(D^*)$
 - A3 Calcular a distância $d(T, T^*)$
 - A4 Aceitar θ se $d < \epsilon$, então voltar a A1.
-

Os parâmetros do algoritmo ABC são o estatístico T , a distância $d(\cdot, \cdot) \geq 0$ e o nível de tolerância $\epsilon > 0$. A aproximação da distribuição a posteriori proporcionada pelo

Algoritmo A é obtida como a marginal de θ da distribuição conjunta

$$\pi_\epsilon(\theta, D^*|D) = \frac{\pi(\theta)f(D^*|\theta)I_{A_{\epsilon,D}}(D^*)}{\int_{A_{\epsilon,D} \times \Theta} \pi(\theta)f(D^*|\theta)dD^*d\theta}, \quad (3.8)$$

onde $I_B(\cdot)$ denota a função indicadora do conjunto B e

$$A_{\epsilon,D} = \{D \in \mathcal{D} : d(T, T^*) \leq \epsilon\}.$$

Em (VOSS, 2013) é estabelecido que se T é um estatístico suficiente e além disso $\epsilon \rightarrow 0$, então a distribuição de $\pi(\theta|d(T, T^*) < \epsilon) \rightarrow \pi(\theta|T)$, que é equivalente com a distribuição de $\pi(\theta|D)$. Em outras palavras, a justificativa para a abordagem ABC é que, quando usamos um estatístico suficiente T e uma tolerância pequena ϵ , temos

$$\pi_\epsilon(\theta|D) = \int \pi_\epsilon(\theta, D^*|D)dD^* \approx \pi(\theta|D),$$

a distribuição a posteriori $\pi(\theta|D)$ é igual ao limite quando ϵ tende a zero de $\pi_\epsilon(\theta|D)$. Dado que ϵ não pode ser o valor zero, (MARIN et al., 2012) sugere considerar ϵ como um pequeno percentil das distâncias simuladas no passo 3 do Algoritmo A.

Como o algoritmo A é computacionalmente ineficiente, (MARJORAM et al., 2003) propõe uma versão mais eficiente que pode ser resumida da seguinte forma:

Algoritmo 2 – ABC-MCMC.

-
- B0 Use o Algoritmo A para obter uma realização de θ^0
 - B1 Gerar $\theta^{(t)}$ a partir do kernel de Markov $q(\cdot|\theta^{(t-1)})$
 - B2 Gerar D^* usando o modelo $f(\cdot|\theta^t)$, então calcule $T^* = T(D^*)$
 - B3 Se $d(T, T^*) < \epsilon$, ir para B4, caso contrário, ficar em $\theta^{(t-1)}$, então voltar a B1.
 - B4 Calcular $h = h(\theta^{(t)}, \theta^{(t-1)}) = \min \left(1, \frac{\pi(\theta^{(t)})q(\theta^{(t-1)}|\theta^{(t)})}{\pi(\theta^{(t-1)})q(\theta^{(t)}|\theta^{(t-1)})} \right)$
 - B5 Aceitar $\theta^{(t)}$ com probabilidade h , caso contrário, ficar em $\theta^{(t-1)}$, então voltar a B1.
-

Uma característica importante do Algoritmo B é que a taxa de aceitação está diretamente relacionada ao valor da verdadeira função de verossimilhança, dado o vetor de parâmetros proposto (SISSON; FAN; TANAKA, 2007), por outro lado, para um valor de ϵ pequeno, é obtida uma baixa taxa de aceitação que afeta a cadeia de Markov, que em situações particulares pode ser presa em regiões de baixa densidade de probabilidade a posteriori, (MARJORAM et al., 2003). Este problema ocorre em regiões de baixa probabilidade, dada a proposta $\theta^{(t)}$, é muito improvável que sejam produzidos dados simulados D^* próximos dos dados observados D , em tal situação, a probabilidade

de que a cadeia se mova fora da região de baixa probabilidade é próxima de zero, desta forma o algoritmo produz amostras altamente dependentes.

Embora todas as cadeias de MCMC correm o risco de ficar presas, o algoritmo ABC-MCMC é suscetível a este tipo de problemas devido aos dois critérios que devem ser cumpridos para aceitar uma proposta $\theta^{(t)}$, primeiro, a probabilidade de aceitação das amostras Metropolis-Hastings primeiro, a probabilidade de aceptación de las muestras Metropolis-Hastings, e segundo, a geração de dados suficientemente próximos aos dados observados, Portanto, a taxa de rejeição do ABC-MCMC pode ser extraordinariamente alta, por isso requer ciclos computacionais bastante altos, mesmo para problemas relativamente simples, (TURNER; ZANDT, 2012).

O Algoritmo ABC Controle de Rejeição Parcial (ABC-PRC por suas siglas em Inglês) proposto por (SISSON; FAN; TANAKA, 2007) como uma opção aos problemas que apresenta o Algoritmo ABC-MCMC, e foi o primeiro algoritmo ABC que usa um filtro de partículas.

O ABC-PRC exige que escolhamos um kernel de transição $K_t(\cdot|\theta^*)$ como uma função de densidade. Tamanho da amostra efetivo (ESS por suas siglas em Inglês) calcula o número equivalente de amostras aleatórias necessárias para obter uma estimativa, de modo que sua variação de Monte Carlo seja igual àquela das N partículas ponderadas. E pode ser resumido da seguinte forma:

Dependendo da natureza do modelo, provar a suficiência de um estatístico T pode ser analiticamente impraticável. Por esta razão, devemos usar o conhecimento sobre o modelo para mostrar que o estatístico T conserva grande parte da informação disponível sobre os parâmetros. (JOYCE; MARJORAM et al., 2008) mostra que isso pode ser alcançado selecionando o estatístico T que minimiza a entropia a posteriori.

3.3 Suficiência

A escolha apropriada de estatísticos de resumo para o uso de métodos ABC corresponde a um estágio fundamental, dado que uma má escolha, ou a seleção de um estatístico insuficiente, poderia significar perda de informação e, em última instância, que o método seja impreciso. Na prática, são raros os casos em que se tem acesso a estatísticos suficiente de forma analítica, por isso, devemos nos preocupar por aqueles estatísticos que podem ser mais representativos, ou seja, aqueles que têm uma perda mínima de informação em relação aos parâmetros de interesse.

A seguir, são apresentados elementos básicos sobre Entropia e informação com o propósito de avaliar o grau de informação dos estatísticos propostos, para mais detalhes veja (COVER; THOMAS, 2012).

Algoritmo 3 – ABC-PRC .

-
-
- C1 Inicializar $\epsilon_1, \dots, \epsilon_T$, e especifique a distribuição de amostragem inicial μ_1 .
Definir indicador de população $t = 1$.
- C2 Definir indicador de partículas $i = 1$.
- C2.1 Se $t = 1$, amostra $\theta^{**} \sim \mu_1(\theta)$ independentemente de μ_1 .
Se $t > 1$, amostra θ^* a partir de a população anterior $\{\theta_{t-1}^{(i)}\}$
com pesos $\{W_{t-1}^{(i)}\}$, e perturbar a partícula para $\theta^{**} \sim K_t(\theta|\theta^*)$
de acordo com um kernel K_t .
Gere um conjunto de dados $D^{**} \sim f(D|\theta^{**})$.
Se $d(T, T^*) \geq \epsilon_t$, então ir para C2.1.
- C2.2 Set
- $$\theta_t^{(i)} = \theta^{**} \text{ e } W_t^{(i)} = \begin{cases} \pi(\theta_t^{(i)})/\mu_1(\theta_t^{(i)}) & \text{if } t = 1 \\ \frac{\pi(\theta_t^{(i)})}{\sum_{j=1}^N W_{t-1}^{(j)} K_t(\theta_t^{(i)}|\theta^*)} & \text{if } t \geq 1 \end{cases}$$
- onde $\pi(\theta)$ denota a distribuição a priori para θ , e
Se $i < N$, aumentar $i = i + 1$ e ir para C2.1.
- C3 Normalize os pesos para que $\sum_{i=1}^N W_t^{(i)} = 1$.
Se $ESS = [\sum_{i=1}^N (W_t^{(i)})^2]^{-1} < E$ depois reamostrar com substituição,
as partículas $\{\theta_t^{(i)}\}$ com pesos $\{W_t^{(i)}\}$ para obter uma nova população
 $\{\theta_t^{(i)}\}$, e definir pesos $\{W_t^{(i)} = 1/N\}$
- C4 Se $t < T$, aumentar $t = t + 1$ e ir para C2.
-
- Onde W_t são os pesos dada a nova partícula θ^{**} .

3.3.1 Entropia e Informação Mútua.

Definição 10. A Entropia de X , denotada por H , mede a incerteza de X e é definida da seguinte forma,

$$H(X) = - \sum_x p_X(x) \log p_X(x) = -\mathbb{E}_X (\log p_X(X)) = -\mathbb{E}_X \left(\frac{1}{\log p_X(X)} \right) \geq 0,$$

onde \mathbb{E}_X denota a esperança sob a função de massa de probabilidade p_X . Sejam (X, Y) variáveis aleatórias discretas com f.d. conjunta $p_{X,Y}$. A entropia condicional $H(Y|X)$ é definida como

$$H(Y|X) = -\mathbb{E}_{X,Y} [\log p_{Y|X}(Y|X)].$$

A informação mútua $I(X; Y)$ entre duas variáveis aleatórias discretas X e Y mede a quantidade de informações que Y contém sobre X . mide la cantidad de información que Y contiene sobre X . Pode-se ver a redução da incerteza sobre X devido ao conhecimento

de Y ,

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ &= \sum_{x,y \in \mathcal{X}} p_{X,Y}(x,y) \log \left(\frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} \right) \\ &= KL(p_{X,Y} || p_X p_Y) \geq 0, \end{aligned}$$

onde $KL(P||Q)$ refere-se à divergência *Kullback-Leibler* (KL) entre duas medidas de probabilidade P y Q . A informação mútua $I(X; Y)$ é igual a 0 se e somente se as variáveis aleatórias X e Y são independentes.

A informação mútua condicional de variáveis aleatórias discretas X , Y e Z é definida como:

$$I(X; Y|Z) = H(X|Z) - H(X|Y, Z);$$

isto é a redução da incerteza de X devido ao conhecimento de Y quando Z é dado. Tal valor é de 0 se e somente se X e Y são condicionalmente independentes, dado Z , o que significa que Z contém todas a informação sobre X em Y . A informação mútua condicional satisfaz a regra da cadeia; para variáveis aleatórias discretas X_1, X_2, \dots, X_n e Y , quer dizer

$$I(X_1, X_2, \dots, X_n; Y) = \sum_{i=1}^n I(X_i; Y|X_1, \dots, X_{i-2}, X_{i-1}).$$

Neste trabalho, é considerada a entropia, em vez de entropia diferencial, com aplicações a variáveis aleatórias contínuas.

3.3.2 Desigualdade do Processamento de Dados e Estatísticos Suficientes

A *desigualdade do processamento de dados* (DPE) estado que para variáveis aleatórias X , Y e Z tal que $X \rightarrow Y \rightarrow Z$, (isto é, Y depende, aleatória ou deterministicamente, de X e Z depende de Y)

$$I(X; Y) \geq I(X; Z),$$

com igualdade somente se $X \rightarrow Y \rightarrow Z$ formam uma cadeia de Markov, que significa que a variável aleatória X e Z são condicionalmente independentes dado Y , isto é, $p_{X,Z|Y} = p_{X|Y}p_{Z|Y}$.

Consideramos agora a família de distribuições $\{f(\cdot|\theta)\}_{\theta \in \Theta}$ e seja X uma amostra de uma distribuição nesta família. Seja \mathcal{S} um estatístico e denotamos por S a variável aleatória tal que $S = \mathcal{S}(X)$, então $\theta \rightarrow X \rightarrow S$. Obtendo por DPE,

$$I(\theta; S) \leq I(\theta; X).$$

Diremos que o estatístico \mathcal{S} é suficiente para o parâmetro θ se e somente se $I(\theta; S) \leq I(\theta; X)$.

Proposição 1. S é um estatístico suficiente para o parâmetro θ se e somente se,

$$I(\theta; X|S) = 0.$$

Neste caso,

$$\mathbb{E}_{\theta, X} \left[\log \frac{p(\theta|X)}{p(\theta|S)} \right] = 0.$$

Exemplo

Como ilustração, consideraremos a estimação do parâmetro de localização μ de uma distribuição de Cauchy. Para isto, foram realizadas simulações com a seguinte configuração de parâmetros: $\mu = 0$ e $\sigma = 1$. Como se trata do parâmetro de localização, um estatístico de resumo possível para o Algoritmo C é a média da amostra. Especificamos como a priori de μ a distribuição *Uniforme* em $(-10, 10)$, e o kernel correspondente $K_t(\mu_t|\mu_{t-1})$ como um passeio aleatório Gaussiano. O Algoritmo C foi executado para 1000 iterações. As tolerâncias programadas são dadas na Tabela 1.

Iteraciones t	Valor de ϵ_t
1-100	10
101-200	5
201-300	2
301-400	1
401-500	0.5
501-600	0.2
601-700	0.1
701-800	0.05
801-900	0.02
901-1000	0.01

Tabela 1 – Simulações e níveis de tolerância.

Para verificar o desempenho do Algoritmo 3, foi considerado como um estimador a média a posteriori, e o desvio padrão da distribuição a posteriori. Os resultados são apresentados na Tabela 2, onde é possível ver uma diferença considerável entre as estimativas pontuais do parâmetro, associadas ao tamanho de ϵ , também se pode ver uma redução significativa em no desvio padrão.

Os resultados mostrados na Tabela 2 e a Figura 5 sugerem que uma boa maneira de ver o Algoritmo C de (SISSON; FAN; TANAKA, 2007) é como uma série de sucessivas

Valor de ϵ	sample mean	(SE)
10	-1.22680	(4.6367)
5	-0.55045	(3.5321)
1	0.12922	(2.7633)
2	0.06258	(2.2970)
0.5	-0.2387	(5.4334)
0.2	0.29349	(2.1673)
0.1	-0.13251	(2.1652)
0.05	0.23571	(1.8058)
0.02	0.36734	(1.5159)
0.01	0.29067	(1.4938)

Tabela 2 – Média e desvio padrão para distribuição a posteriori, dependendo das tolerâncias e assumindo que $\mu = 0$.

aplicações do Algoritmo A sobre variáveis aleatórias extraídas de uma distribuição a priori, que é uma convolução do núcleo suavizado e as variáveis selecionadas da rodada anterior de rejeição.

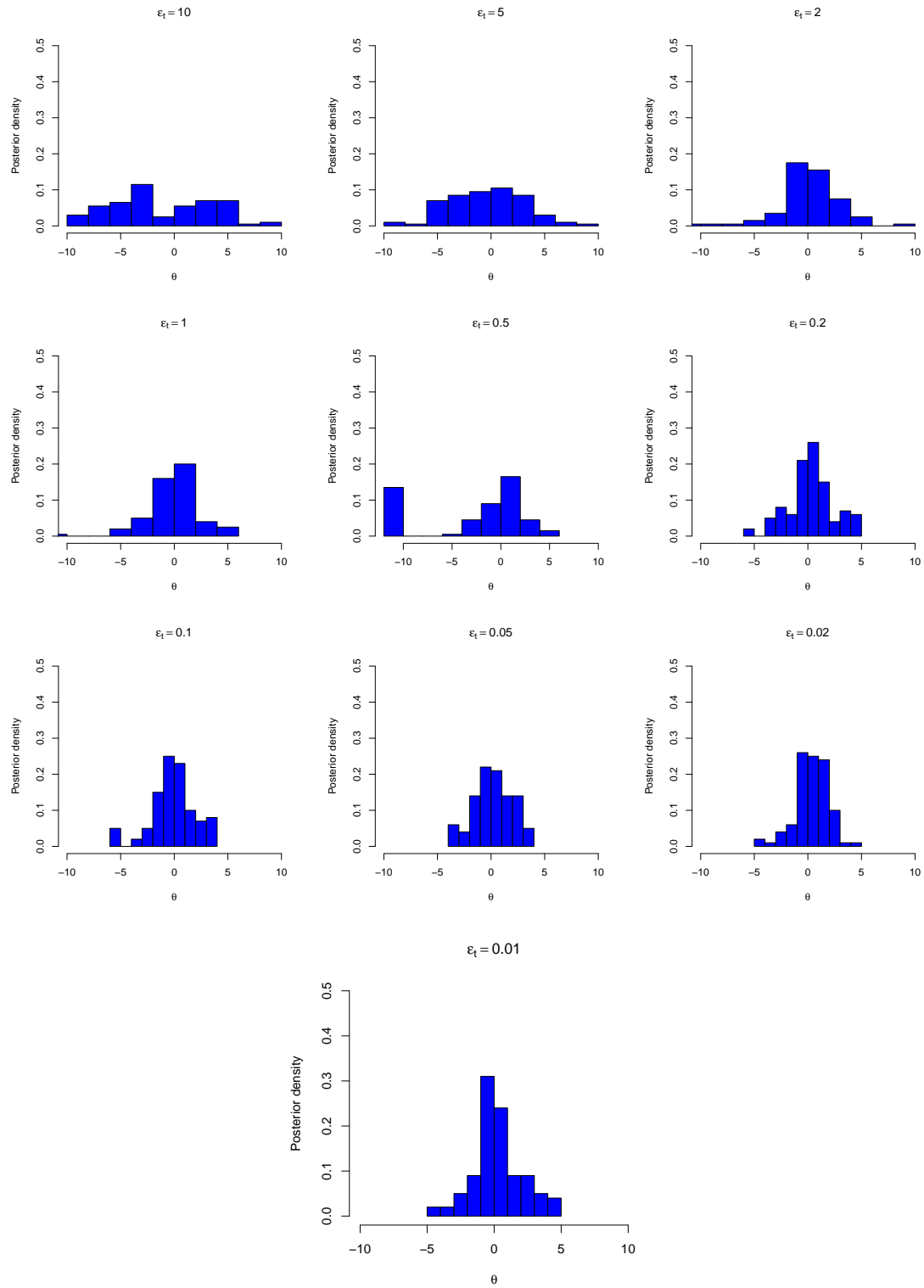


Figura 5 – A posteriori, dependendo do vetor de tolerâncias ϵ .

Parte IV

Resultados obtidos

4 Estimadores Bayesianos para o sBm

4.0.1 Resumo

Em alguns problemas que envolvem séries temporais de variáveis aleatórias contínuas, considerar barreiras semipermeáveis permite construir modelos mais precisos para o fenômeno estudado. Por exemplo, em finanças, é possível considerar barreiras psicológicas semipermeáveis que envolvem preços de ativos, gerando diferentes procedimentos de tomada de decisão em cada lado da barreira. Desta forma, propomos um modelo que capta essa característica em termos de tempos locais, dando origem ao chamado movimento Browniano assimétrico (sBm).

Recentemente, vários trabalhos consideraram o sBm em modelagem ou em problemas de simulação, assim como em alguns problemas de otimização. Ver, por exemplo (LEJAY, 2006) para referências sobre o tema, e para um estudo sobre as diversas construções e possíveis aplicações do sBm.

Neste trabalho estamos interessados em estimar o parâmetro de assimetria quando observamos uma trajetória do processo em uma quadrícula de tempo igualmente espaçados.

Do ponto de vista estatístico, este problema representa uma análise intermédio entre o problema clássico da estimativa do drift em uma difusão, (KUTOYANTS, 1984; LIPSTER; SHIRYAEV, 2001), e a estimativa da variância de uma difusão, ver (FLORENS-ZMIROU, 1993; JACOD, 1998) e referências no mesmo, onde as medidas de probabilidade geradas pelas trajetórias são singulares para diferentes valores do parâmetro.

Até onde sabemos, existem poucos trabalhos sobre a estimação deste parâmetro. Em (BARDOU; MARTINEZ, 2010), os autores assumem que sBm é refletido nos níveis 0 e 1 para garantir a ergodicidade, por outro lado em (LEJAY; MORDECKI; TORRES, 2014), o estimador de máxima verossimilhança é construído e é analisado seu comportamento assintótico sob a hipótese nula do parâmetro de assimetria nulo, ou seja, sBm sendo o movimento browniano padrão. Sob hipóteses gerais, no entanto, o problema não é tratado analiticamente e ainda não foi estudado

Neste trabalho, discute-se uma análise baseada em simulação sobre o comportamento assintótico dos estimadores de Bayes para o parâmetro de assimetria, mostrando sua consistência através de simulações. A metodologia é aplicada a uma trajetória que os leões marinhos da América do Sul (SASL, *Otaria flavescens*) seguem durante suas viagens de alimentação

La metodología se aplica a una trayectoria que los lobos marinos sudamericanos

(SASL, *Otaria flavescens*) siguen durante sus viajes de alimentación, tais viagens no mar alternam-se com visitas a colônias para descansar e que as fêmeas amamentam filhotes de lactação ((BONESS; BOWEN; OFTEDALL, 1994)), na costa do sul do Chile

4.0.2 Formulação do modelo

Ao considerar o sBm X com parâmetro $\theta \in [-1, 1]$ definido em 2.1, para $t \in [0, T]$ e o esquema de amostragem denotado por $X_i := X_{i\Delta}$, $i = 0, \dots, n$, e $\Delta = T/n$. Como é mostrado em (WALSH, 1978). As transições entre estados, mostradas na Figura 6, são definidas pelo kernel

$$q(t, x, y|\sigma, \theta) = (\phi(y - x|t\sigma^2) + \mathbf{sgn}(x, y)\theta\phi(|x| + |y|t\sigma^2))\frac{1}{2}(2 - \mathbf{1}_{\{0\}}(x)), \quad (4.1)$$

onde

$$\phi(x|t\sigma^2) = \frac{1}{\sqrt{2\pi}t\sigma} \exp\left(-\frac{x^2}{2t\sigma^2}\right),$$

é a densidade de uma variável aleatória Gaussiana centrada com variância $t\sigma^2$, $\mathbf{sgn}(x, y) = \mathbf{sgn}(y)\mathbf{1}_{\{y \neq 0\}}(y) + \mathbf{sgn}(x)\mathbf{1}_{\{y = 0\}}(y)$, e $\mathbf{sgn}(x) = 1$ se $x > 0$, $\mathbf{sgn}(x) = -1$ se $x < 0$ e $\mathbf{sgn}(x) = 0$ se $x = 0$. Além disso, podemos reescrever as transições entre os estados como:

$$q(t, x, y|\sigma, \theta) = \begin{cases} \phi(y - x|t\sigma^2) + \theta\phi(x + y|t\sigma^2) & \text{si } y > 0 \\ \phi(y - x|t\sigma^2)(1 - \theta) & \text{c.o.c.} \end{cases}, \quad (4.2)$$

quando $x > 0$,

$$q(t, x, y|\sigma, \theta) = \begin{cases} \phi(y - x|t\sigma^2) - \theta\phi(x + y|t\sigma^2) & \text{si } y \leq 0 \\ \phi(y - x|t\sigma^2)(1 + \theta) & \text{c.o.c.} \end{cases}, \quad (4.3)$$

quando $x < 0$, e por ultimo quando $x = 0$,

$$q(t, 0, y|\sigma, \theta) = \begin{cases} \frac{1}{2}\phi(y|t\sigma^2) + (1 - \theta) & \text{si } y \leq 0 \\ \frac{1}{2}\phi(y|t\sigma^2) + (1 + \theta) & \text{c.o.c.} \end{cases}. \quad (4.4)$$

Pode-se observar desde o kernel de transição, que o parâmetro θ tem uma grande influência na trajetória somente quando está perto da origem.

tiene una gran influencia sobre la trayectoria sólo cuando está cerca del origen (barreira de semipermeabilidade). De fato, para um valor fixo do tempo t , e quando x está o suficientemente “longe” do origem, o segundo termo é quase zero. Nesse caso, a trajetória é praticamente um movimento browniano, quase sem influência de θ , e, portanto, quase não existe informações sobre θ . Por outro lado, quando uma trajetória se aproxima ao origem desde o primeiro quadrante, por exemplo, se $\theta < 0$ a trajetória é “empurrada” para

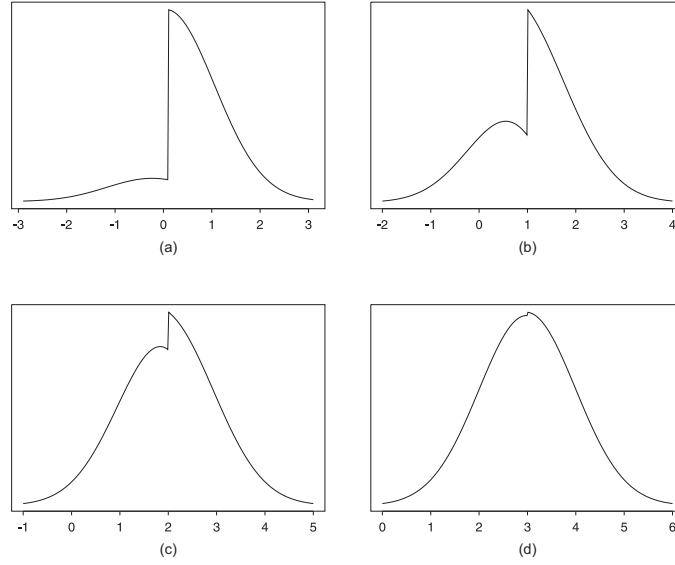


Figura 6 – Kernel de transição $q(t, x, y | \tau, \theta)$, com $t = 1$, $\tau = 1$, considerando $\theta = 0.8$ e: (a) $x = 0.1$, (b) $x = 1$, (c) $x = 2$, (d) $x = 3$.

cruzar ao quarto quadrante, e se $\theta > 0$, tem mais probabilidade de refletir-se e permanecer no primeiro quadrante.

No caso limite $\theta = 1$, a solução para 2.1 é o movimento browniano **refletido** sobre o origem, e quando $\theta = -1$, é o movimento browniano **refletido** baixo o origem, tão logo a trajetória se torna negativa. O caso $\theta = 0$ corresponde ao movimento browniano padrão. A Figura 12 (a) mostra uma trajetória típica para $\theta = 0.8$.

4.0.3 Tempo de Saída

Consideramos $\tau_x = \inf\{t > 0 : X_t = x\}$, como o primeiro tempo de visita ao ponto $x \in \mathbb{R}$. Sea $J = (a, b)$ qualquer intervalo aberto tal que a probabilidade de saída (exiting probability) $P_x(\tau_{\{a,b\}} < \infty)$ é igual a 1, para qualquer $x \in J$. Então existe uma função contínua, estritamente crescente $S(x)$ em \mathbb{R} , tal que

$$P_x(\tau_a < \tau_b) = \frac{S(b) - S(x)}{S(b) - S(a)}, \quad x \in J.$$

A função S é chamada função de escala do processo. Para mais detalhes sobre funções de escala, veja por exemplo (BREIMAN, 1992). Para o movimento browniano assimétrico, a função de escala é dada por,

$$S(x) = \begin{cases} \alpha^{-1}x & \text{si } x \geq 0 \\ (1 - \alpha)^{-1}x & \text{c.o.c.} \end{cases}$$

onde $\alpha = (1 + \theta)/2$, como em (LEJAY, 2006). Isso implica que, se $a = -b < 0 < b$,

$$P_0(\tau_{-b} < \tau_b) = 1 - \alpha = (1 - \theta)/2.$$

Podemos ver que esse valor não depende de b , somente de θ :

- si $\theta = 0$, então $P_0(\tau_{-b} < \tau_b) = 1/2$;
- si $\theta > 0$, então $P_0(\tau_{-b} < \tau_b) < 1/2$;
- si $\theta < 0$, então $P_0(\tau_{-b} < \tau_b) > 1/2$,

para qualquer $b > 0$. Em outras palavras, se $\theta < 0$, uma vez que a partícula atinge a barreira, ela é empurrada para baixo com probabilidade $(1 - \alpha)$ maior a $1/2$; e se $\theta > 0$, a partícula é empurrada para cima com probabilidade α maior a $1/2$.

Do ponto de vista dos procedimentos inferenciais, dada uma trajetória observada, para aquelas observações o suficientemente perto da barreira, as proporções destas em cada lado da barreira são funções lineares de θ . Esta propriedade pode nos guiar para definir uma distribuição a priori apropriada para este parâmetro.

4.0.4 Verossimilhança e configurações a priori

Pelo exposto, podemos ver que para o processo inferencial, após observar uma trajetória, se esta “passa” longe do origem, quase não dá informação sobre θ ; pelo contrário, quanto maior seja o número de passos o suficientemente perto da barreira, mais informativos são os dados sobre θ . Este ganho de informação será clara em na forma analítica da densidade a posteriori para θ

Para o modelo 2.1, vamos considerar o espaço paramétrico

$$\Omega = \{\boldsymbol{\omega} = (\theta, \sigma^2) \in \mathbb{R}^2 : \sigma^2 > 0, -1 \leq \theta \leq 1\}.$$

Para uma amostra dada $\mathbf{d} = (x_1, \dots, x_n)$ do processo observado aos tempos (t_1, \dots, t_n) , a função de verossimilhança para $\boldsymbol{\omega} \in \Omega$, depois de observar \mathbf{d} é dada por

$$l_{\mathbf{d}}(\boldsymbol{\omega}) = \prod_{i=1}^{n-1} q(\tau_i, x_i, x_{i+1} \mid \theta, \sigma^2),$$

com q como em (4.1), y $\tau_i = t_{i+1} - t_i$. Para facilitar a notação, seja $a_i = a_i(\sigma^2) = \phi_{\tau_i \sigma^2}(x_{i+1} - x_i)$ y $b_i = b_i(\sigma^2) = \phi_{\tau_i \sigma^2}(x_i + x_{i+1})$. Com esta notação, a função de

verossimilhança de (θ, σ^2) é proporcional a

$$\begin{aligned}
 l_{\mathbf{d}}(\boldsymbol{\omega}) &\propto \left(\prod_i a_i(\sigma^2) \right) (1 + \theta)^{n_{-+}} (1 - \theta)^{n_{+-}} \prod_{\substack{x_j > 0 \\ x_{j+1} > 0}} \left(1 + \frac{b_j(\sigma^2)}{a_j(\sigma^2)} \theta \right) \times \\
 &\quad \times \prod_{\substack{x_j < 0 \\ x_{j+1} < 0}} \left(1 - \frac{b_j(\sigma^2)}{a_j(\sigma^2)} \theta \right) \\
 &\propto \frac{1}{\sqrt{\sigma^{2n}}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_i (x_{i+1} - x_i)^2 \right\} (1 + \theta)^{n_{-+}} (1 - \theta)^{n_{+-}} \times \\
 &\quad \times \prod_j \left(\frac{a_j(\sigma^2)}{b_j(\sigma^2)} \pm \theta \right),
 \end{aligned} \tag{4.5}$$

onde n_{-+} é o número de cruzamentos do lado negativo para o positivo, n_{+-} é o número de cruzamentos do lado positivo para o lado negativo, e $b_j/a_j = \exp\{-2x_j x_{j+1}/(\tau_j \sigma^2)\}$.

Dada uma trajetória, quando x_i está distante da barreira, e $|x_i| + |x_{i+1}|$ fica grande, menos informação temos sobre θ , como é intuitivo que ocorra. Assim, a parte mais relevante da amostra, para θ , são aquelas seções próximas da barreira.

Observe que, para cada passo que atravessa a barreira, se $x_i \leq 0$ e $x_{i+1} > 0$, então $q(\tau_i, x_i, x_{i+1} \mid \theta, \sigma^2) \propto (1 + \theta) a_i$. Por outro lado, se $x_i \geq 0$ e $x_{i+1} < 0$, então $q(\tau_i, x_i, x_{i+1} \mid \theta, \sigma^2) \propto (1 - \theta) a_i$. No primeiro caso, $\theta = -1$ torna-se uma raiz da função de verossimilhança, e quando ocorre o segundo, $\theta = 1$ torna-se raiz.

Por outro lado, para cada dois passos consecutivos positivos ou negativos, $q(\tau_j, x_j, x_{j+1} \mid \theta, \sigma^2)$, como uma função linear em θ , tem uma raiz, $\pm a_j/b_j$, fora do intervalo $[-1, 1]$, com máximo em $\theta = 1$ se $x_j > 0$, e em $\theta = -1$, no outro caso. Isto implica, por exemplo, que se observarmos uma trajetória positiva sem cruzamentos com a barreira, a estimativa de máxima verossimilhança para θ é 1, gerando evidencia de un movimiento Browniano reflejado, (LEJAY; MORDECKI; TORRES, 2014, Lemma 1).

Suponhamos que temos informação a priori sobre assimetria e volatilidade, e denotamos por $f(\boldsymbol{\omega})$ a densidade a priori para $\boldsymbol{\omega}$, quantificando essa informação. Pelo teorema de Bayes, a densidade a posteriori para $\boldsymbol{\omega}$, dada a amostra \mathbf{d} , é determinada pela relação 3.2.

Se assumirmos uma simetria a priori ao redor de $\theta = 0$, a simetria a posteriori para θ basicamente depende dos factores polinomiais definidos por os passos realizados em cada um dos lados da barreira. Se, por exemplo, uma trajetória com pelo menos duas cruces tem mais passos no lado positivo do que no lado negativo, então a distribuição a posteriori está inclinada para a esquerda, dando mais massa de probabilidade para $\theta > 0$.

Além disso, quando a trajetória não atravessa a barreira, se $x_i > 0$, para tudo i , então a distribuição a posteriori é assimétrica à direita, com valor modal em $\theta = 1$, pelo

contrário, se $x_i < 0$, para tudo i , então a distribuição a posteriori é assimétrica à esquerda, com valor modal em $\theta = -1$.

Consideramos a priori que, θ e σ^2 são variáveis aleatórias independentes. Para o inverso da variância adotamos uma distribuição Gama, $G(\alpha_s, \beta_s)$, isso devido principalmente ao primeiro termo na expressão de verossimilhança (4.5), desta forma, podemos obter um tipo de conjugação para $1/\sigma^2$. Por outro lado, a partir da interpretação de θ como uma função linear de uma proporção dada na Seção 4.0.3 e novamente para obter uma conjugação parcial para θ , adotamos para θ uma distribuição a priori Beta em $[-1, 1]$ com hiperparâmetros α_t, β_t .

Com isso, a densidade a priori é dada por

$$f(\theta, \sigma^2) \propto (1 + \theta)^{\alpha_t - 1} (1 - \theta)^{\beta_t - 1} \left(\frac{1}{\sigma^2} \right)^{\alpha_s + 1} \exp \left(-\frac{\beta_s}{\sigma^2} \right) \mathbf{1}_{\Omega}(\theta, \sigma^2),$$

e conseqüentemente, a densidade a posteriori pode ser escrita da seguinte forma

$$\begin{aligned} f(\theta, \sigma^2 \mid \mathbf{d}) &\propto \left(\frac{1}{\sigma^2} \right)^{\alpha_s + 1 + n/2} \exp \left\{ -\frac{1}{\sigma^2} \left(\beta_s + \sum_i \frac{1}{2\tau_i} (x_{i+1} - x_i)^2 \right) \right\} \times \\ &\times (1 + \theta)^{n_{-+}} (1 - \theta)^{n_{+-}} \prod_{\substack{x_j > 0 \\ x_{j+1} > 0}} \left(1 + \frac{b_j(\sigma^2)}{a_j(\sigma^2)} \theta \right) \times \\ &\times \prod_{\substack{x_j < 0 \\ x_{j+1} < 0}} \left(1 - \frac{b_j(\sigma^2)}{a_j(\sigma^2)} \theta \right). \end{aligned}$$

Os dois últimos fatores mostram a dependência entre θ e σ^2 , dada pela informação actualizada depois de observar \mathbf{d} , quanto mais forte a dependência, mais a trajetória permanece próxima da barreira. À medida que a trajetória se afasta da barreira, esses fatores convergem exponencialmente para um.

4.0.5 Considerações sobre o comportamento assintótico de a posteriori.

Nesta seção apresentamos uma análise baseada em simulações sobre o comportamento assintótico da distribuição a posteriori de θ , dado σ^2 , na medida que o tamanho da amostra aumenta.

Em (LEJAY; MORDECKI; TORRES, 2014), foi alcançado um primeiro passo em direção a um resultado de convergência para o estimador de máxima verossimilhança (MLE). Os autores caracterizam a distribuição limite sob a hipótese nula do movimento Browniano padrão, quer dizer $\theta = 0$, ea convergência está relacionada com \mathcal{G} -estabilidade (ver (JACOD, 1998) para a definição e propriedades). Este tipo de convergência é comprovada para uma vizinhança de $\theta = 0$, mas mesmo nesse caso não é uma condição suficiente

para a consistência do estimador. Nesse sentido, (LEJAY; MORDECKI; TORRES, 2014) é o único resultado analítico sobre a convergência do estimador para o parâmetro de assimetria θ em o sBm.

Além disso, para o MLE, a estimativa obtida para uma sequência de tempos $\{iT/n\}_i$, é equivalente à estimativa obtida para uma sequência de passos de tempo igual a um. Assim, as simulações realizadas consideram o aumento dos tempos $\{1, \dots, n\}$ e fornecem suporte para a consistência da distribuição posterior quando o tamanho da amostra n aumenta, como podemos ver a seguir.

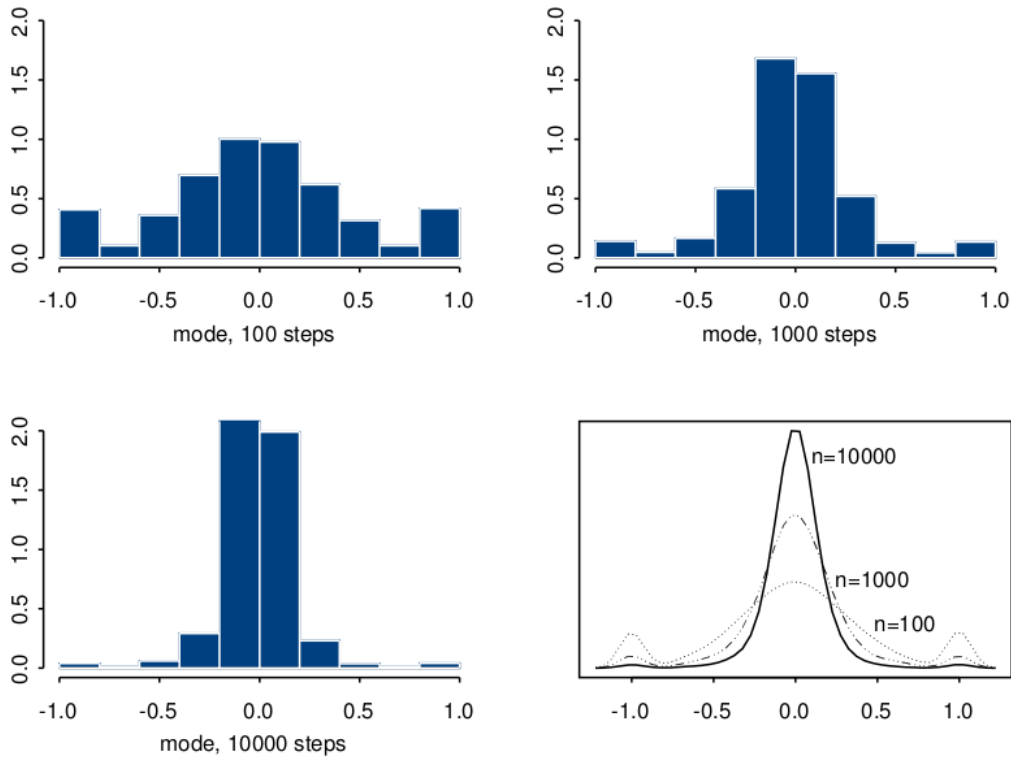


Figura 7 – Distribuição amostral da moda a posteriori dadas trajetórias com n passos, $n = 100, 1000, 10000$ assumindo $\theta = 0$.

Para esta discussão, apresentamos a distribuição da amostra simulada da moda a posteriori, a média a posteriori, e os quartis a posteriori 0.025 e 0.975 para θ . Estas distribuições foram obtidas simulando dez mil trajetórias com n passos, considerando $n = 100, 1000$ e 10000 , para diferentes valores de θ em -1 a 1 . Consideramos para o desenvolvimento das simulações, o valor de $\sigma^2 = 1$; resultados análogos foram obtidos para outros valores de σ^2 .

A Figura 7 apresenta a distribuição amostral da moda a posteriori de θ , a partir de trajetórias geradas com $\theta = 0$.

Em A Figura 7 observa-se também que para trajetórias com $n = 100$ passos, a distribuição amostral da moda a posteriori é trimodal, com valores modais $-1, 0, 1$. Essas modas extremas ocorrem porque existe uma probabilidade positiva de que a trajetória não atinja a origem, dando evidência para $\theta = 1$ se a trajetória é principalmente positiva, e para $\theta = -1$, se fosse principalmente negativo. Por simetria, ambos casos ocorrem com a mesma probabilidade, como mostrado pelas simulações. Quando n aumenta, essa probabilidade tende a zero, como podemos ver nos histogramas, para $n = 1000$ e $n = 10000$ passos, onde a massa de probabilidade em torno de -1 e 1 quase desaparece.

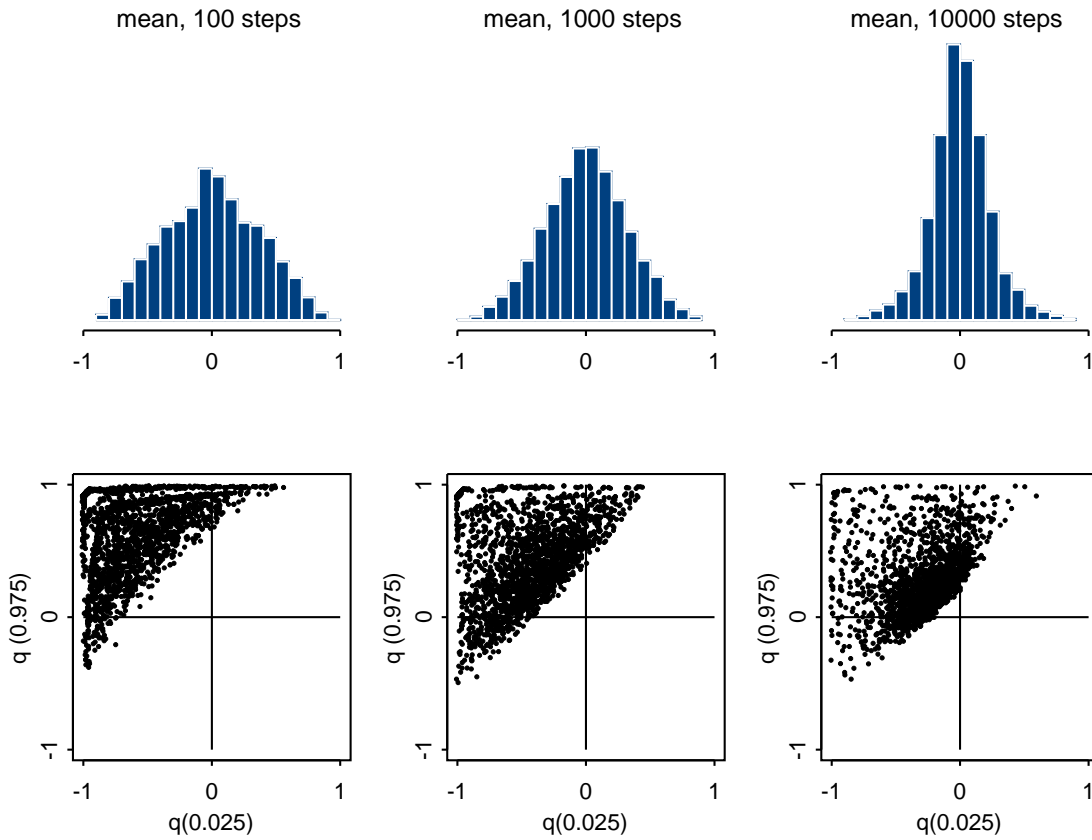


Figura 8 – Média a posteriori e quartis $q(0.025)$ y $q(0.975)$, para trajetórias com n passos, $n = 100, 1000, 10000$, assumindo $\theta = 0$.

A Figura 8 apresenta a distribuição da amostra da média a posteriori e a distribuição da amostra conjunta para os quantis $q(0.025)$ e $q(0.975)$ a posteriori. É claro que a distribuição da amostra da média para a posteriori está concentrada em torno de $\theta = 0$ quando n cresce. Além disso, a distância $q(0.975) - q(0.025)$ diminui a medida que n aumenta, mostrando que a distribuição a posteriori se concentra assintoticamente. A Tabela 3 mostra a média da amostra e o desvio padrão para a média e a posteriori.

A convergência também pode ser observada na distribuição amostral da média a posteriori para diferentes valores de θ , como é mostrado na Figura 9. Nos gráficos, o

	$n = 100$		$n = 1.000$		$n = 10.000$	
	média	(DP)	média	(DP)	média	(DP)
moda	0.000783	(0.494)	-0.00256	(0.319)	-0.00192	(0.192)
média	0.00565	(0.366)	-0.000689	(0.312)	-0.000387	(0.226)

Tabela 3 – Média e desvio padrão para moda e média a posteriori, assumindo que $\theta = 0$.

eixo horizontal representa os valores nominais para θ , desde -1 até 1. Para cada θ , as linhas sólidas representam intervalos de credibilidade do 95% para a média a posteriori, dependendo do número de cruzamentos. Em (a) com 100, (b) 1000 e (c) 10000 passos, a linha contínua considera todas as trajetórias, a linha pontilhada mostra o mesmo intervalo para a média a posteriori, mas considerando apenas as trajetórias que têm cinco cruces ou mais, e a linha segmentada, aqueles com dez cruces ou mais com a barreira de semipermeabilidade. O intervalo de credibilidade, para a média a posteriori, é mais preciso quando o número de cruzamentos com a barreira aumenta, e, nesse sentido, é necessário enfatizar que o número de cruzamentos é uma estatística relevante para ter uma inferência mais precisa sobre θ .

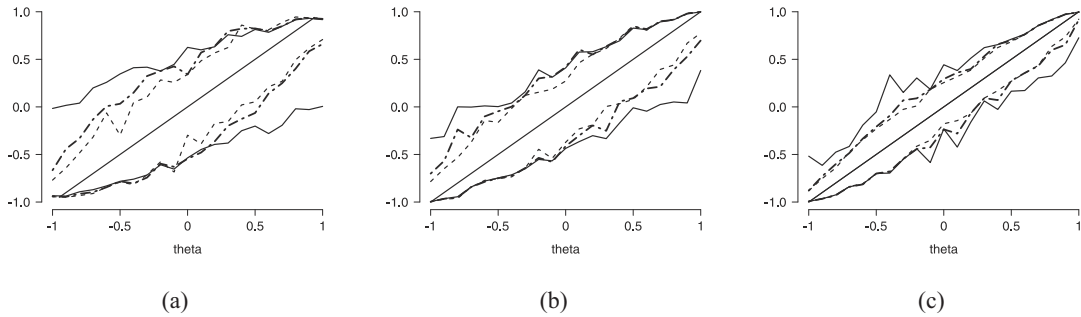


Figura 9 – Intervalo de amostragem para a média a posteriori de θ , com $\sigma^2 = 1$, considerando $\theta \in [-1, 1]$ e: (a) $n = 100$, (b) $n = 1000$, (c) $n = 10000$ passos. Para cada gráfico, a linha sólida representa o 95% dos valores centrais da média a posteriori, a linha pontilhada mostra o mesmo intervalo considerando apenas as trajetórias que possuem cinco cruces ou mais, e a linha segmentada, aquelas com dez cruces ou mais.

Podemos ver na Figura 10, que a proporção de trajetórias positivas, sem atravessar a barreira, é maior para trajetórias com $n = 100$ passos, como é evidenciado pela moda a posteriori perto de 1. No entanto, essa proporção diminui à medida que n aumenta.

A Figura 11 se apresenta a distribuição da amostra da média a posteriori, dependendo do número de cruces com a barreira semi-permeável. É apreciada uma diminuição na variabilidade da distribuição a posteriori à medida que o número de cruzamentos aumenta.

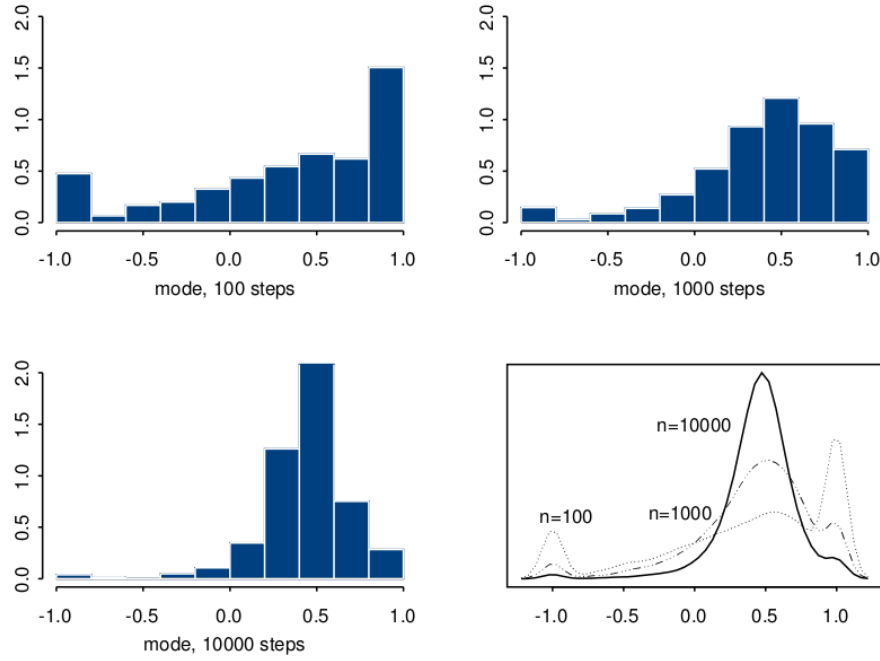


Figura 10 – Distribuição da amostra da moda a posteriori, dadas as trajetórias de n passos, $n = 100, 1000, 10000$ assumindo $\theta = .5$.

Como pode ser visto nas Figuras 7, 10 e 8, para tamanhos de amostra inferiores a 1000, a moda a posteriori para o parâmetro de assimetria mostra uma distribuição amostral mais dispersa do que a média a posteriori. Isto significa que, em certo sentido, a média a posteriori é um estimador mais confiável para θ . Por outro lado, ambos estimadores melhoram à medida que o número de cruzamentos com a barreira aumenta, e nesse sentido a sua consistência é bem suportada por simulações, para qualquer valor de θ .

4.1 Computação para dados

4.1.1 Dados simulados

Esta seção se descreve o procedimento seguido para obter inferências a posteriori, a partir de uma única trajetória, as estimativas pontuais a posteriori e e -valor. Para fins ilustrativos, considerarmos a trajetória simulada com $n = 1000$ passos, a partir de um sBm com parâmetros $\sigma^2 = 1$ y $\theta = 0.8$, tal trajetória é apresentada na Figura 12(a).

A distribuição a priori considerada para $1/\sigma^2$ é Gama $[1, 1]$ e para θ quatro distribuições a priori Beta diferentes, com parâmetros (α_t, β_t) igual a $(1, 1)$, $(0.5, 0.5)$, $(5, 1)$ e $(1, 5)$. Desde que a distribuição a posteriori $f(\omega \mid \mathbf{d})$ tem uma forma analítica fechada, exceto pela constante de normalização, a calculamos em uma grade com 141×40 pontos

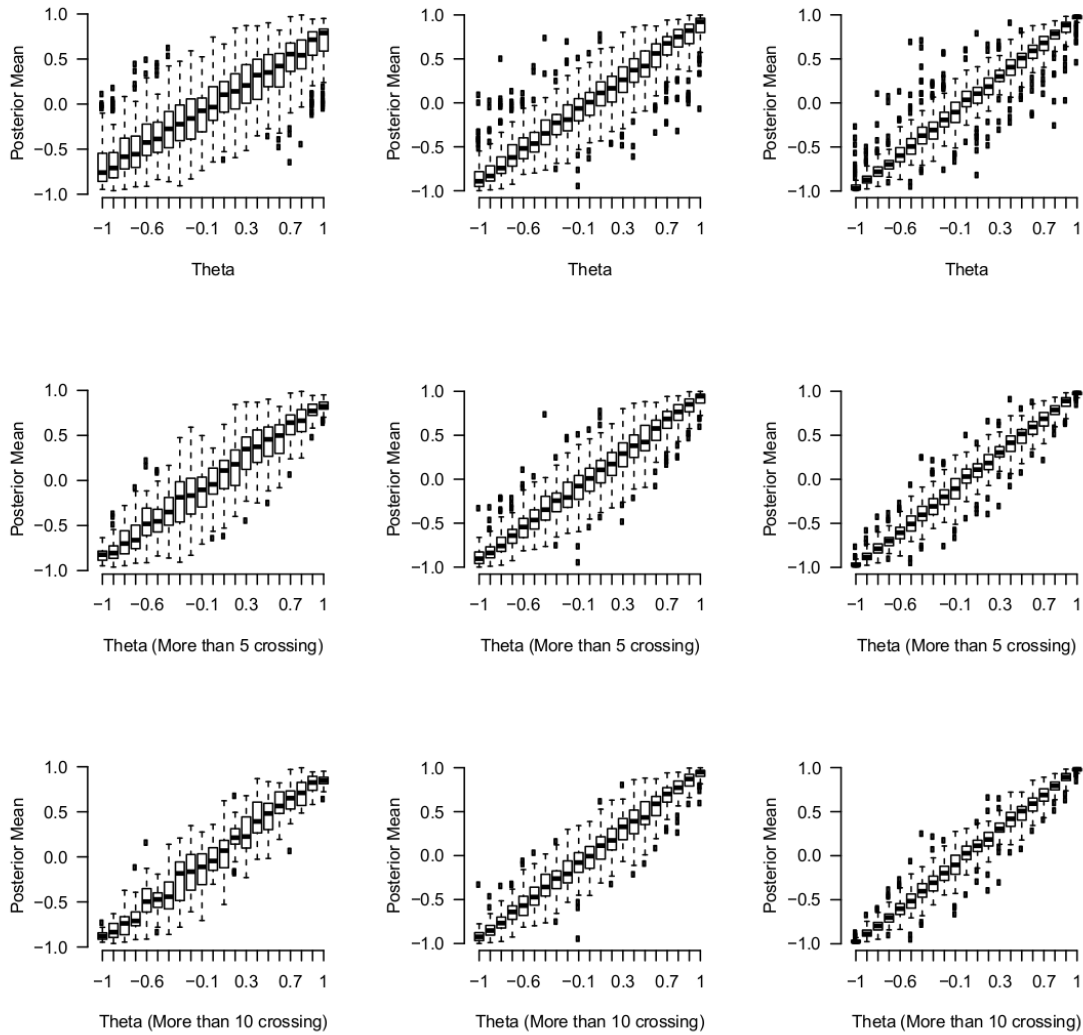


Figura 11 – Boxplot da distribuição amostral da média a posteriori, considerando o número de cruzamentos com a barreira.

para (θ, σ^2) em $[-1, 1] \times [0.8, 1.2]$. As distribuições a posteriori correspondentes a cada uma das a priori consideradas, são representadas na figura 12 em azul, rosa, ciano e verde, respectivamente..

É possível observar que praticamente não existe diferença entre as distribuições a posteriori obtidas nas três primeiras a priori. Esta falta de influência deve-se à grande quantidade de passos próximos da barreira de semipermeabilidade, que fornecem, como já mencionado, informação relevante sobre o parâmetro de assimetria θ . Com a versão da distribuição a priori Beta(1,5), a estimativa pontual é fortemente próxima a $\theta = -1$, neste caso, a diferença nas estimativas para θ não passa despercebida, como podemos ver na Tabela 4. Mesmo neste último caso, o e -valor em favor da hipótese $\theta = 0$ é insignificante, o que também concorda com o fator Bayes.

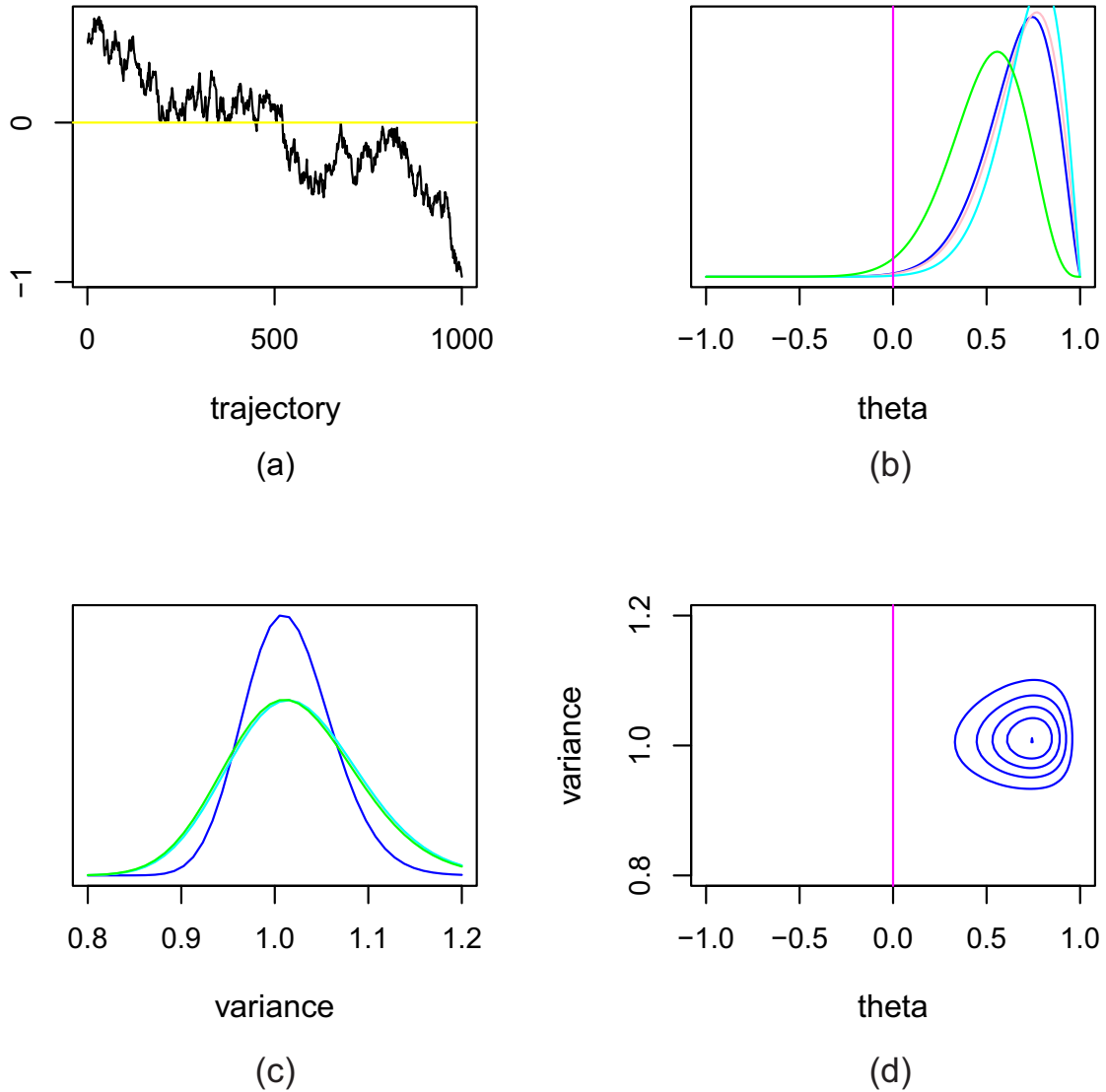


Figura 12 – (a) Simulação de uma trajetória do processo, com $n = 1000$ passos, para $\theta = 0.8$ e $\sigma^2 = 1$. Note que a trajetória tende a ser mais positiva que negativa. (b) Densidades a posteriori marginais e conjuntas para (θ, σ^2) , com distribuição a priori para θ : Beta(1,1) em azul, Beta(0.5,0.5) em rosa, Beta(5,1) em ciano, Beta(1,5) em verde. A linha vertical magenta representa a hipótese $\theta = 0$.

4.1.2 Real data

Sete SASL (1 macho e 6 fêmeas) foram capturados em julho de 2009 e junho de 2010 na costa em frente a Calbuco ($41^{\circ}48'S$; $73^{\circ}08'W$), sul do Chile. Os indivíduos foram marcados com etiquetas GPS da Unidade de Investigação em Mamíferos Marinhos (SRDL, por suas siglas em inglês), colados na pele dorsal. As etiquetas SRDL-GPS coletaram dados sobre a posição dos animais, o comportamento de mergulho e a temperatura da água. Estes dados foram processados a bordo para calcular uma série de parâmetros, depois resumidos e transmitidos através do sistema ARGOS.

Neste trabalho, modelamos os locais de viagem de alimentação de um desses

distribuição a priori para θ	moda a posteriori para (θ, σ^2)	média a posteriori para (θ, σ^2)	e -valor para $\theta = 0$	fator de Bayes contra $\theta = 0$
Beta(1,1)	(0.74, 1.01)	(0.66, 1.02)	0.0029	16.8
Beta(1,5)	(0.56, 1.02)	(0.49, 1.02)	0.0224	2.55

Tabela 4 – Estimativas a posteriori para (θ, σ^2) , considerando diferentes distribuições a priori para θ , tomando uma trajetória simulada com valores nominais $\theta = 0.8$ e $\sigma^2 = 1$. As duas últimas colunas mostram a evidência em favor da hipótese $\theta = 0$, dadas por o e -valor e o fator de Bayes.

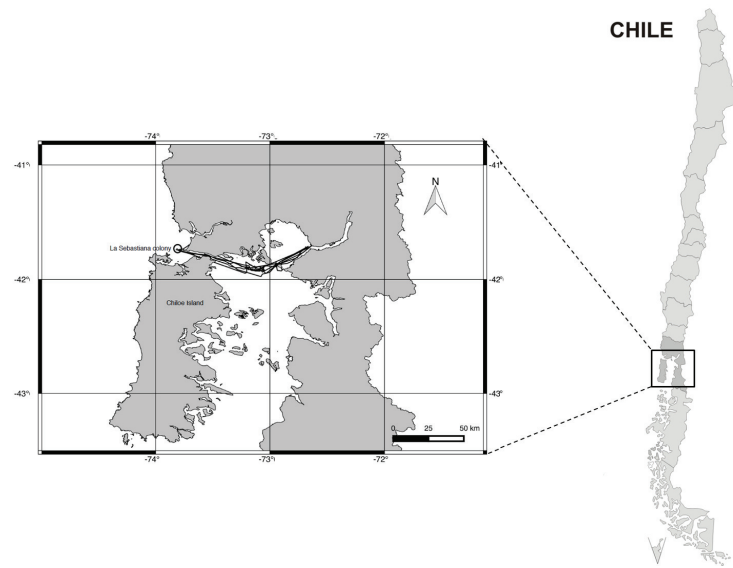


Figura 13 – Localização da área de alimentação, Calbuco, Chile.

SASL (sujeito ID 96721), como exemplo de trajetória assimétrica. Vários estudos sobre diferentes espécies de leões marinhos mostraram uma fidelidade às áreas de alimentação entre indivíduos (veja por exemplo, (BONADONNA et al., 2001; ROBSON et al., 2004; CALL et al., 2008)), incluído SASL na costa do Atlântico ((RIET-SAPRIZA et al., 2013; RODRÍGUEZ et al., 2013)). A fidelidade individual às áreas de alimentação pode permitir aos animais localizar locais produtivos de alimentação em viagens sucessivas, o que poderia ser uma estratégia benéfica, uma vez que confere a vantagem de reduzir os custos totais de viagem ((ROBSON et al., 2004; RIET-SAPRIZA et al., 2013)). Com base nisso, levantamos a hipótese de que o SASL viajará sempre para as mesmas áreas de alimentação e, portanto, dentro do modelo proposto, 0 deve estar próximo de -1 .

A maioria dos viagens de alimentação estão concentradas entre a colônia de La Sebastiana e as águas interiores da ilha de Chiloé (Figura 13). Como é mostrado na Figura 14(a) existe uma barreira física na latitude -41.76 que corresponde à localização da colônia de leões marinhos.

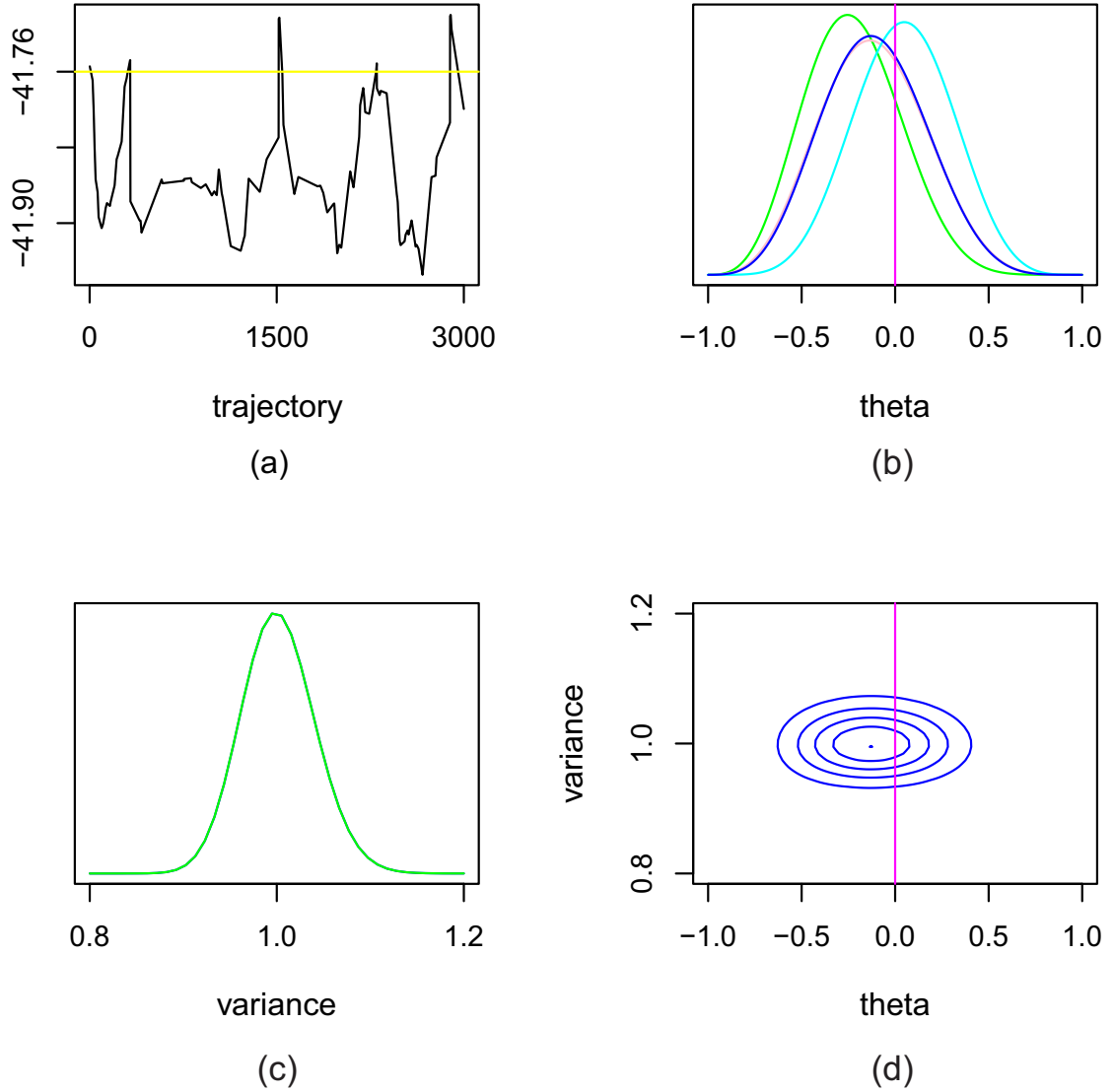


Figura 14 – (a) Latitude de uma trajetória de um leão-marinho na costa de frente a Calbuco, no sul do Chile. (b) Densidades a posteriori marginal e conjunta para o parâmetro de assimetria θ , com distribuição a priori: Beta(1,1) em azul, Beta(0.5,0.5) em rosa, Beta(5,1) em ciano, Beta(1,5) em verde. A linha vertical magenta representa a hipótese $\theta = 0$.

Como no exemplo simulado, consideramos quatro distribuições a priori diferentes para θ , $\text{Beta}(\alpha_t, \beta_t)$, com (α_t, β_t) igual a (1, 1), (0.5, 0.5), (5, 1) e (1, 5), representado na Figura 14 em azul, rosa, ciano e verde, respectivamente. Também padronizamos a trajetória observada, dividindo todos os termos pelo desvio padrão das diferenças observadas e a distribuição a priori considerada para $1/\sigma^2$ é Gama(1,1). As estimativas a posteriori são apresentadas na Tabela 5.

A evidência dada pelo fator de Bayes contra H é favorável a esta hipótese. No entanto, considerando a regra de decisão (3.6) e uma distribuição a priori uniforme para θ ,

a priori para θ	moda a posteriori para (θ, σ^2)	média a posteriori para (θ, σ^2)
Beta(1,1)	(−0.13, 0.99)	(−0.11, 1.00)
Beta(5,1)	(0.04, 0.99)	(0.04, 1.00)
Beta(1,5)	(−0.26, 0.99)	(−0.23, 1.00)

a priori para θ	e-valor para $\theta = 0$	fator de Bayes contra $\theta = 0$	$P(\theta < 0 \mid \mathbf{d})$
Beta(1,1)	0.656	0.394	0.648
Beta(5,1)	0.876	0.283	0.424
Beta(1,5)	0.355	0.404	0.799

Tabela 5 – Estimativas a posteriori para (θ, σ^2) , considerando diferentes distribuições a priori para θ , para a trajetória SAL observada. As últimas três colunas mostram a evidência em favor da hipótese $\theta = 0$, dada pelo e -valor, o fator de Bayes e a probabilidade a posteriori de $\theta < 0$.

devemos tomar a decisão de rejeitar H se e somente se,

$$0.656(w_0 + c) < w_1 + c \iff \frac{2}{3}w_0 < w_1 + \frac{1}{3}c, \text{ aproximadamente.}$$

Em outras palavras, se temos, por exemplo $w_1 \approx c$, e o custo do erro de tipo II é pelo menos a metade do custo do erro tipo I, então H deve ser rejeitada.

Além disso, as probabilidades $P(\theta < 0 \mid \mathbf{d}) > 0.4$, para as diferentes distribuições a priori, indicam que é provável que o sujeito nem sempre viaje para a mesma área de alimentação, mas em vez disso algumas de suas viagens foram para latitudes mais baixas. Biologicamente, pelo menos três explicações diferentes, não excludentes mutuamente, poderia explicar por que o SASL se move para diferentes áreas de alimentação. Primeiro, as presas podem mudar de localização devido a variações nas características oceanográficas, forçando os predadores a mudar suas áreas de alimentação. (FRIEDLAENDER et al., 2006). Segundo, os indivíduos geralmente segregam e mudam seus locais de alimentação para diminuir a competição por recursos ((LEUNG et al., 2012)). Finalmente, SASL é considerado um predador generalista, o que significa que consome diferentes espécies como presa, (CAPPOZZO; PERRIN,). Portanto, se a abundância e a disponibilidade de presas diminuem em uma determinada área de alimentação, o indivíduo pode apresentar plasticidade comportamental deixando a área para se alimentar em um lugar diferente ou procurar presas diferentes ((SIGLER et al., 2009)).

5 Estimadores bayesianos para el FATGBM

A estimativa dos parâmetros da distribuição do tipo G é uma tarefa difícil, (BROWNE; MCNICHOLAS, 2015), esta dificuldade está em na forma da função de verossimilhança com relação aos parâmetros. Está documentado que os métodos de otimização padrão não resolvem o problema de inferência, em particular ao estimar todos os parâmetros do modelo (BARNDORFF-NIELSEN; BLAESILD, 1981). Tem havido consideráveis contribuições a este respeito, por exemplo, (BLAESILD; SØRENSEN, 1992) propõe o algoritmo “Hyp” para estimar os parâmetros apenas para o caso $\lambda = 1$, no entanto, o método tem um alto custo computacional, (EBERLEIN; PRAUSE, 2002). Por outro lado (PROTASSOV, 2004) usa o algoritmo EM e realiza a estimativa dos parâmetros apenas para o caso $\lambda = -1/2$. A estimação dos parâmetros sob a abordagem Bayesiana, de acordo com a bibliografia vista até o momento, foi feita apenas na distribuição NIG (Normal Inverse Gaussian), (LILLESTØL, 2000) , (KARLIS; LILLESTÖL, 2004).

Neste capítulo estamos interessados na implementação do ABC-PMC, para os modelos apresentados na seção 2.4, considerando um subordinado τ_t distribuído Gama com parâmetros α_0 e β_0 , ambos conhecidos. Os análises desenvolvidos são baseadas em um conjunto de simulações do modelo.

5.1 Formulação do Modelo

O processo FATGBM com parâmetro $\omega = (\theta, \sigma) \in \mathbb{R} \times \mathbb{R}^+$ definido em 2.11, para tudo $t \in [0, T]$, tem densidade marginal definida em 2.15, e representada nas Figuras 15 e 16 para diferentes configurações de parâmetros.

5.1.1 Simulações

Existem vários métodos de simulação de um processo estocástico tipo G, neste trabalho as simulações serão realizadas através de processos estocásticos com incrementos reais e independentes, tendo em conta que também depende da medida de probabilidade do processo T_t dada na equação 2.11.

O problema subjacente é que o vetor T não é observado, mas como a distribuição do processo não depende de ω , e os parâmetros dos quais depende são conhecidos $\alpha = 0.5 = \beta$, é possível realizar simulações do processo T , e obter resultados dos métodos ABC proposto. No entanto, o estimador não é fixo e, obviamente, depende do vetor simulado T .

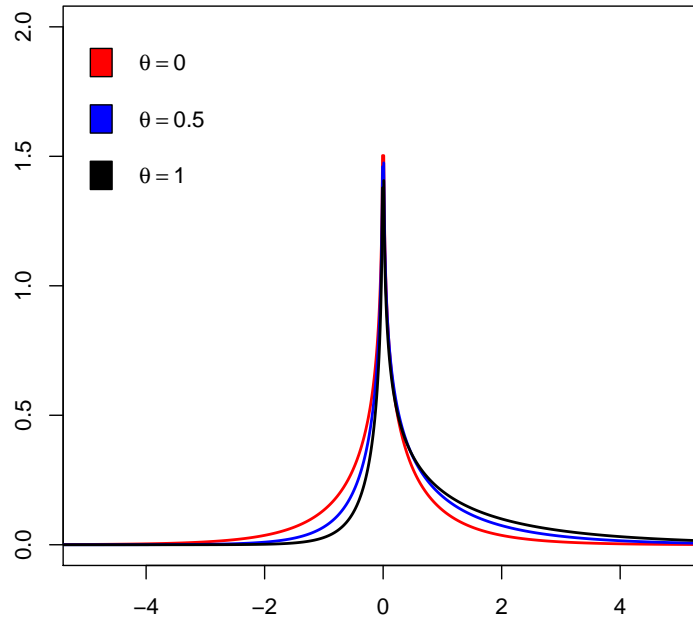


Figura 15 – O kernel de transição, variando o parâmetro de assimetria θ , com $t = 1$ e $\sigma = 1$.

5.1.2 Configuração de distribuição a priori

A implementação do Algoritmo C exige a abordagem de três questões importantes: a seleção das distribuições a priori para os parâmetros envolvidos no modelo, o kernel de transição associado aos pesos e os estatísticos de resumo.

As distribuições a priori foram consideradas independentes, isto é,

$$\pi(\theta, \sigma^2) \propto \pi_1(\theta) \cdot \pi_2(\sigma^2), \quad (5.1)$$

onde $\pi_1(\cdot)$ é uma distribuição normal com hiperparâmetros $\mu_0 = 0$, $\sigma_0^2 = 2$ e $\pi_2(\cdot)$ é uma distribuição gama com hiperparâmetros $\alpha_0 = 2$ e $\beta_0 = 3$. Em relação ao kernel de transição para o método ABC-PMC, $K_t(\cdot|\gamma)$ será considerado como a densidade Normal com parâmetros γ e 0.1. Além disso, foi estudado o desempenho de 6 combinações de estatísticos de resumo que, a priori, têm informações sobre os parâmetros do modelo. As medidas de resumo são apresentadas na Tabela 6. A seleção é a combinação que gera os melhores resultados ao combinar o Algoritmo A e os critérios descritos na seção 3.3.

Para estudar o comportamento do estimador bayesiano proposto para o modelo 2.10, foram realizadas simulações com diferentes configurações de parâmetros θ e σ . O parâmetro de localização μ é fixado em US 0 e não será estimado. Um total de três

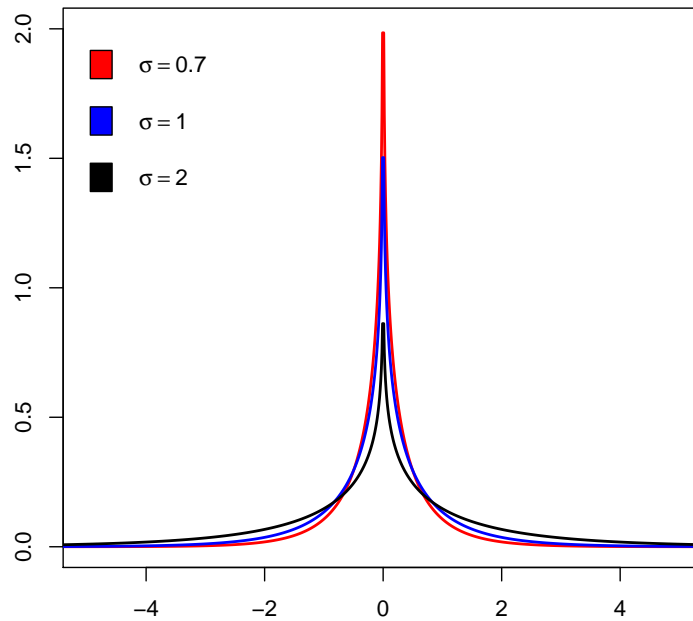


Figura 16 – O kernel de transição, variando o parâmetro de escala σ , com $t = 1$ e $\theta = 0$.

θ	σ
Terceiro momento	Coefficiente de variação
Coefficiente de assimetria de Fisher	Desvio padrão
Coefficiente de assimetria de Pearson	

Tabela 6 – Estatísticos usados para avaliar a suficiência.

especificações para os parâmetros foram investigados. As combinações são identificadas como modelos na Tabela 7.

	Modelo 1	Modelo 2	Modelo 3
θ	-.5	1	.5
σ	1	1	1

Tabela 7 – Configuração de parâmetros do modelo.

Para avaliar as propriedades do Algoritmo C e sua eficiência na estimativa das distribuições a posteriori de θ y σ foram geradas 500 amostras de 1000 observações, de cada modelo, utilizando R ((R Core Team, 2013)). Para todas as réplicas do modelo foi calculada a média e moda a posteriori, o que será denotado por $(\hat{\theta}_{mean}, \hat{\sigma}_{mean})$ e $(\hat{\theta}_{mode}, \hat{\sigma}_{mode})$ respectivamente. Finalmente, para medir o desempenho de ambos os estimadores, usamos:

o desvio padrão (SE), erro absoluto AB e a raiz quadrada do erro quadrado médio, os últimos definidos respectivamente como:

$$AB = \frac{1}{n} \sum_{i=1}^n |\hat{\theta}_i - \theta| \quad \text{y} \quad \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{\theta}_i - \theta)^2}.$$

O valor \sqrt{MSE} fornece alguma informação sobre o viés do método quando é comparado com o desvio padrão (SE).

5.1.3 Resultados

Os resultados do AB e MSE para os estimadores propostos são apresentados na Tabela 8, tomando apenas a distribuição a posteriori gerada na última iteração do Algoritmo C para fazer as comparações.

	Model	M_1	M_2	M_3
$\hat{\theta}_{mean}$	AB	0.1215	0.0603	0.0974
	MSE	0.0274	0.0067	0.0154
$\hat{\theta}_{mode}$	AB	0.1277	0.0671	0.1288
	MSE	0.0292	0.0077	0.0289
$\hat{\sigma}_{mean}$	AB	0.0766	0.0675	0.0736
	MSE	0.0097	0.0073	0.0091
$\hat{\sigma}_{mode}$	AB	0.1137	0.0570	0.1151
	MSE	0.0335	0.0052	0.0344

Tabela 8 – Resultado de AB e MSE para os estimadores média e moda a posteriori.

A partir da Tabela 8 pode-se observar que a média e a moda a posteriori produzem resultados semelhantes para as estimativas θ . No entanto, a média a posteriori produz melhores resultados para a estimação de σ , em relação ao AB e MSE.

As Tabelas 9, 10 e 11 mostram a média (SM) e desvio padrão (EP) da média a posteriori, para cada modelo considerado. O vetor de tolerância considerado é $\epsilon = (2, 1, .5, .1)$.

	$\epsilon = 2$		$\epsilon = 1$		$\epsilon = 0.5$		$\epsilon = 0.1$	
	SM	(SE)	SM	(SE)	SM	(SE)	SM	(SE)
$\hat{\theta}$	-0.4083	(0.182)	-0.5434	(0.130)	-0.6083	(0.122)	-0.5998	(0.083)
$\hat{\sigma}$	0.7261	(0.181)	0.8233	(0.135)	0.8208	(0.096)	0.9650	(0.092)

Tabela 9 – Média e desvio padrão da amostra para a média a posteriori, dependendo do vetor de tolerâncias e assumindo que $\theta = -.5$ e $\sigma^2 = 1$.

	$\epsilon = 2$		$\epsilon = 1$		$\epsilon = 0.5$		$\epsilon = 0.1$	
	SM	(SE)	SM	(SE)	SM	(SE)	SM	(SE)
$\hat{\theta}$	-0.0002	(0.131)	0.0017	(0.119)	-0.0012	(0.112)	0.0006	(0.084)
$\hat{\sigma}$	0.6220	(0.134)	0.5772	(0.114)	0.4961	(0.091)	1.0526	(0.067)

Tabela 10 – Média e desvio padrão da amostra para a média a posteriori, dependendo do vetor de tolerâncias e assumindo que $\theta = 0$ e $\sigma^2 = 1$.

	$\epsilon = 2$		$\epsilon = 1$		$\epsilon = 0.5$		$\epsilon = 0.1$	
	SM	(SE)	SM	(SE)	SM	(SE)	SM	(SE)
$\hat{\theta}$	0.4143	(0.221)	0.5470	(0.164)	0.6012	(0.121)	0.5725	(0.104)
$\hat{\sigma}$	0.9444	(0.181)	0.8223	(0.154)	0.8227	(0.113)	0.9688	(0.091)

Tabela 11 – Média e desvio padrão da amostra para a média a posteriori, dependendo do vetor de tolerâncias e assumindo que $\theta = 0$ e $\sigma^2 = 1$.

	média	(SE)	média	(SE)	média	(SE)
$\hat{\theta}_{mean}$	-0.6016	(0.137)	-0.0004	(0.088)	0.6043	(0.134)
$\hat{\sigma}_{mean}$	0.9116	(0.160)	1.0291	(0.066)	0.9090	(0.161)

Tabela 12 – Média e desvio padrão da amostra para a moda a posteriori, assumindo que $\epsilon = 0.1$ e todas as configurações de parâmetros.

Por espaço e simplicidade, apenas os resultados das simulações consideradas mais relevantes são apresentados em tabelas.

Nas Figuras 17 e 18, as médias a posteriori de $\omega = (\theta, \sigma)$ são incluídas para as réplicas dos Modelos 2 e 3, considerando o vetor de tolerâncias ϵ . É claro a partir deles e das Tabelas 9, 10 e 11 que a variabilidade não é uma constante nas diferentes iterações. De fato, essa variabilidade diminui à medida que ϵ diminui.

A Figura 19 mostra as modas a posteriori conjuntas resultantes de cada réplica dos Modelos 1, 2 e 3, considerando a última iteração do Algoritmo C. Nelas, observa-se maior variabilidade e presença de atípicos, especialmente para o Modelos 1 e 3, resultado que é corroborado pela Tabela 12.

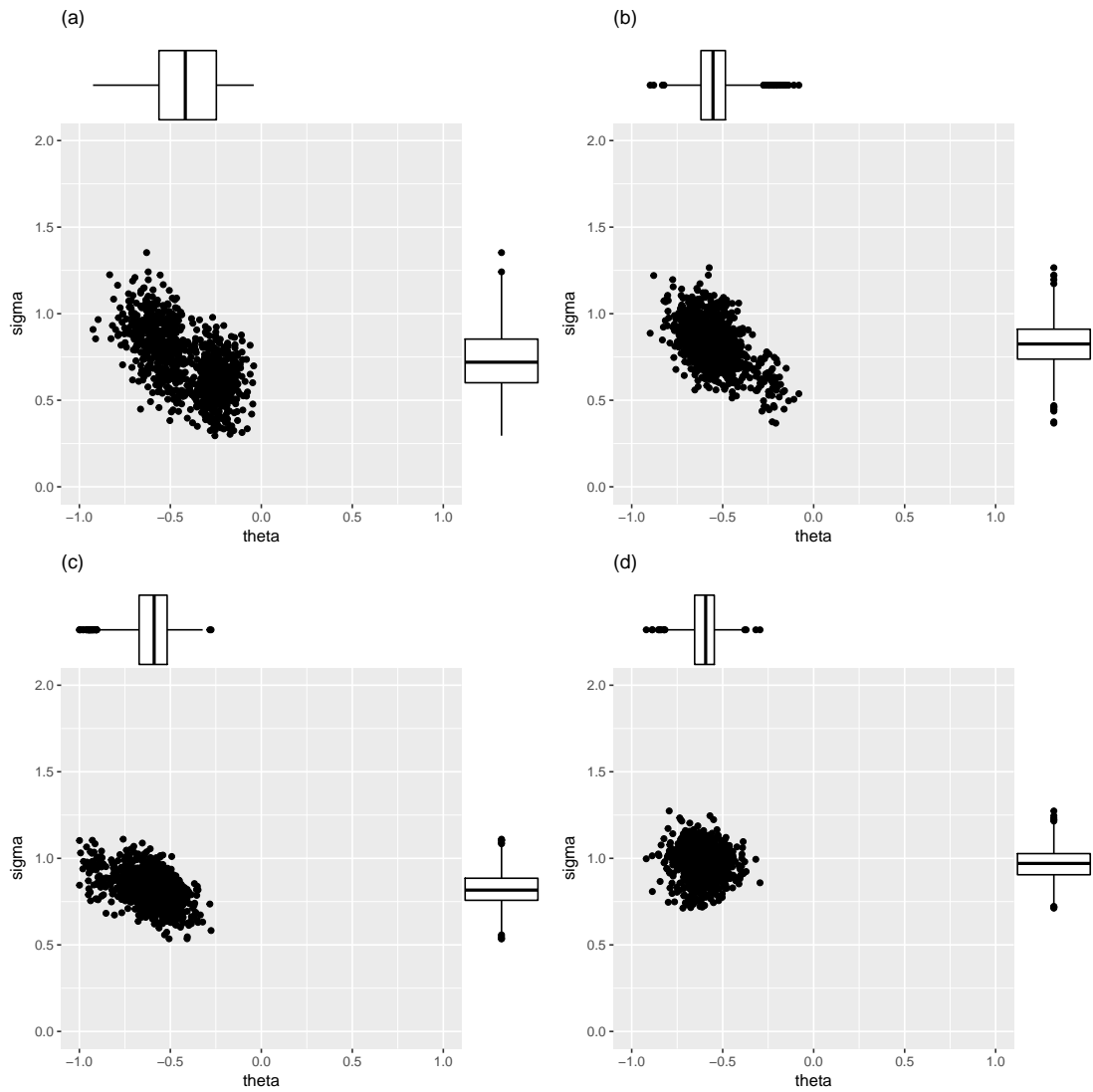


Figura 17 – Média a posteriori conjunta de (θ, σ) , dependendo do vetor de tolerância ϵ e assumindo $\theta = 0$ e $\sigma^2 = 1$

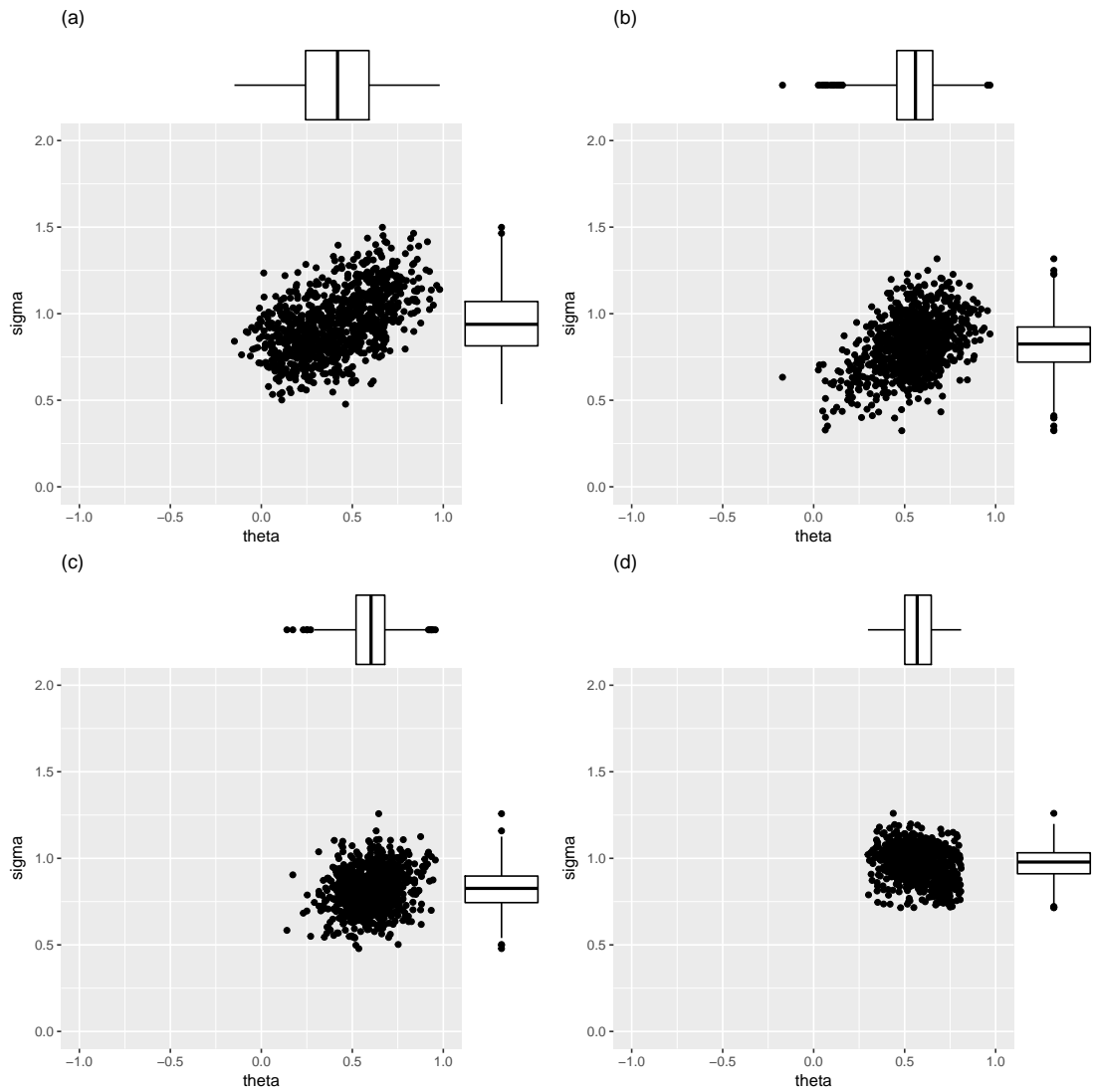


Figura 18 – Média a posteriori conjunta de (θ, σ) , dependendo do vetor de tolerância ϵ e assumindo $\theta = .5$ e $\sigma^2 = 1$

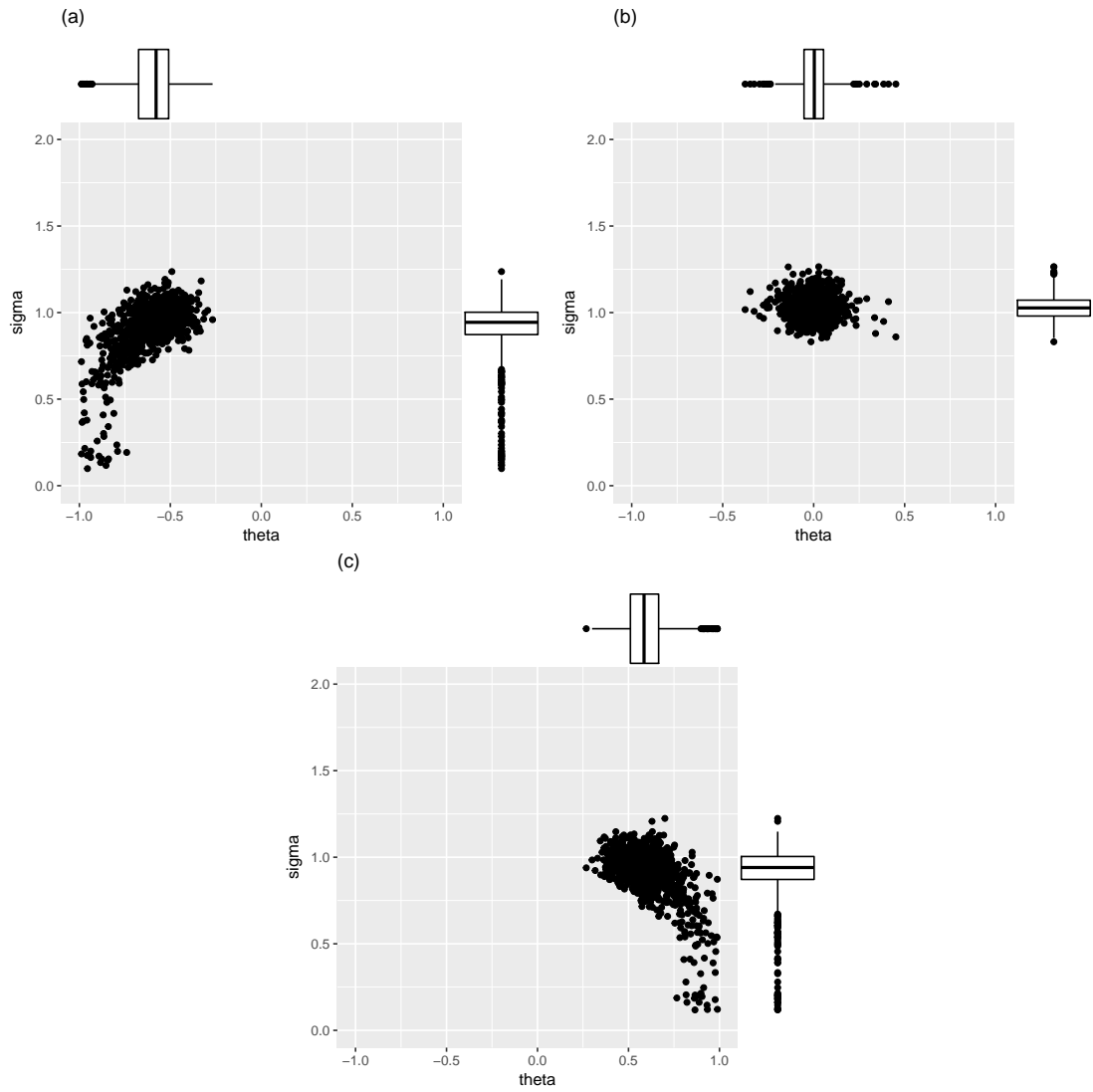


Figura 19 – Moda a posteriori conjunta de (θ, σ) , assumindo que $\epsilon = 0.1$ e todas as configurações de parâmetros.

6 Considerações Finais

6.0.1 Resultados sBm

Em um problema de seleção de modelos, alguns aspectos devem ser levados em conta, como considerações teóricas sobre a problemática em si ou a experiência passada. A amostra observada pode fornecer informações sobre o modelo, principalmente se existirem cruzamentos através da barreira, sin embargo, consideraciones adicionales son necesarias para escoger el sBm como un modelo adecuado para un problema específico.

Do ponto de vista da estimação, como pode ser visto nas Figuras 7 e 10, para amostras com tamanho menor que 1000, a moda a posteriori e, em particular, o estimador de máxima verossimilhança para o parâmetro de assimetria apresentam distribuição amostral com maior dispersão que a média a posteriori.

Isso faz que, em certo sentido, a média a posteriori seja uma estimativa mais confiável para θ . Por outro lado, ambos estimadores são melhoradas quando o número de cruzamentos aumenta na amostra, e nesse sentido a sua consistência é apoiada por meio das simulações.

Uma possível generalização deste problema é considerar que, em cada lado da barreira de semipermeabilidade. Una posible generalización de este problema es considerar que, en cada lado de la barrera de semipermeabilidad, os valores para o parâmetro de volatilidade são diferentes, σ_1^2 e σ_2^2 , permitindo assim diferentes dinâmicas dependendo da barreira de semipermeabilidade. Neste espaço paramétrico expandido, a metodologia proposta também pode ser aplicada. Levando em conta essa parametrização, obtemos uma hipótese precisa que estabelece que essas difusões têm o mesmo valor que o parâmetro $\sigma_1^2 = \sigma_2^2$, cuja evidência pode ser medida pelo e -valor, por exemplo. Outra possível extensão consiste em considerar mais de uma barreira e vários processos de difusão entre elas, ou até mais, considerando uma barreira dependente do tempo, problemas de modelos dinâmicos subjacentes.

A análise proposta aqui poderia ser usada para modelar mudanças comportamentais em predadores, como mamíferos marinhos. Compreender os hábitos de alimentação de animais e como isso varia temporalmente e espacialmente pode ser um indicador de eventos ecossistêmicos, e liga as mudanças na dinâmica dos predadores e presa com a variabilidade ambiental.

6.0.2 Resultados FATGBM

A classe de algoritmos ABC apresentada é ideal no contexto em que a função de verossimilhança é analiticamente intratável. O método proposto parece ser eficiente quando comparado aos principais procedimentos clássicos de estimativa encontrados na literatura. Neste caso, a eficiência foi verificada apenas no contexto de dados simulados e com uma única configuração do processo subordinado. No entanto, isso pode ser facilmente estendido para outras configurações de ruído subordinadas.

Do ponto de vista da estimativa, como pode ser visto nas Figuras 19, a moda a posteriori para o parâmetro variabilidade apresenta distribuição amostral com maior dispersão que as médias a posteriori, principalmente nos Modelos 1 e 3. Isso implica que, em certo sentido, a média a posteriori é um estimador mais confiável para σ .

Uma possível generalização deste problema é considerar que, várias estruturas para o vetor T , e fazer as estimativas dos parâmetros no modelo completo.

Referências

- APPLEBAUM, D. *Lévy processes and stochastic calculus*. [S.l.]: Cambridge university press, 2009. Citado na página 25.
- BACHELIER, L. Theory of speculation. *The random character of stock market prices*, MIT Press, Cambridge, Mass. Blattberg, v. 1018, p. 17–78, 1900. Citado na página 29.
- BARDOU, O.; MARTINEZ, M. Statistical estimation for reflected skew processes. *Stat. Inference Stoch. Process.*, v. 13, p. 231–248, 2010. Citado na página 51.
- BARAHONA, M.; RIFO, L.; SEPÚLVEDA, M.; TORRES, S. A simulation-based study on bayesian estimators for the skew brownian motion. *Entropy*, Multidisciplinary Digital Publishing Institute, v. 18, n. 7, p. 241, 2016. Citado na página 40.
- BARNDORFF-NIELSEN, O. Exponentially decreasing distributions for the logarithm of particle size. *Proc. R. Soc. Lond. A*, The Royal Society, v. 353, n. 1674, p. 401–419, 1977. Citado 2 vezes nas páginas 14 e 31.
- BARNDORFF-NIELSEN, O.; BLAESILD, P. Hyperbolic distributions and ramifications: Contributions to theory and application. In: *Statistical distributions in scientific work*. [S.l.]: Springer, 1981. p. 19–44. Citado na página 66.
- BARNDORFF-NIELSEN, O.; HALGREEN, C. Infinite divisibility of the hyperbolic and generalized inverse gaussian distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, Springer, v. 38, n. 4, p. 309–311, 1977. Citado na página 31.
- BARNDORFF-NIELSEN, O.; NICOLATO, E.; SHEPHARD, N. et al. Some recent developments in stochastic volatility modelling. *Quantitative Finance*, Taylor & Francis, v. 2, n. 1, p. 11–23, 2002. Citado na página 14.
- BEAUMONT, M.; ZHANG, W.; BALDING, D. Approximate bayesian computation in population genetics. *Genetics*, Genetics Soc America, v. 162, n. 4, p. 2025–2035, 2002. Citado 2 vezes nas páginas 14 e 42.
- BERNARDINI, D. de; RIFO, L. Full bayesian significance test for extremal distributions. *Journal of Applied Statistics*, Taylor & Francis, v. 38, n. 4, p. 851–863, 2011. Citado 2 vezes nas páginas 40 e 41.
- BLACK, F.; SCHOLES, M. The pricing of options and corporate liabilities. *Journal of political economy*, The University of Chicago Press, v. 81, n. 3, p. 637–654, 1973. Citado na página 29.
- BLAESILD, P.; SØRENSEN, M. *"Hyp": A Computer Program for Analyzing Data by Means of the Hyperbolic Distribution*. [S.l.]: University of Aarhus, Department of Theoretical Statistics, 1992. Citado na página 66.
- BOCHNER, S. Diffusion equation and stochastic processes. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 35, n. 7, p. 368–370, 1949. Citado na página 26.

- BONADONNA, F.; LEA, M.; DEHORTER, O.; GUINET, C. Foraging ground fidelity and route-choice tactics of a marine predator: the antarctic fur seal *arctocephalus gazella*. *Marine Ecology Progress Series*, v. 223, p. 277–286, 2001. Citado na página 63.
- BONESS, D.; BOWEN, W.; OFTEDALL, O. Evidence of a maternal foraging cycle resembling that of otariid seals in a small phocid, the harbor seal. *Behavioral Ecology and Sociobiology*, v. 34, p. 95–104, 1994. Citado na página 52.
- BRAGANÇA-PEREIRA, C. de; STERN, J. Evidence and credibility: full bayesian significance test for precise hypotheses. *Entropy, Molecular Diversity Preservation International*, v. 1, n. 4, p. 99–110, 1999. Citado na página 40.
- BREIMAN, L. *Probability*. [S.l.]: SIAM, 1992. Citado na página 53.
- BROWNE, R.; MCNICHOLAS, P. A mixture of generalized hyperbolic distributions. *Canadian Journal of Statistics*, Wiley Online Library, v. 43, n. 2, p. 176–198, 2015. Citado na página 66.
- BUCHHOLZ, H. *The confluent hypergeometric function: with special emphasis on its applications*. [S.l.]: Springer Science & Business Media, 2013. v. 15. Citado na página 25.
- CALL, K.; REAM, R.; JOHNSON, D.; STERLING, J.; TOWELL, R. Foraging route tactics and site fidelity of adult female northern fur seal (*callorhinus ursinus*) around the pribilof islands. *Deep Sea Research Part II: Topical Studies in Oceanography*, v. 55, p. 1883–1896, 2008. Citado na página 63.
- CAPPOZZO, H.; PERRIN, W. South american sea lion *otaria flavescens*. In: EDITOR (Ed.). *Encyclopedia of marine mammals*. [S.l.]: Elsevier. Citado na página 65.
- COVER, T.; THOMAS, J. *Elements of information theory*. [S.l.]: John Wiley & Sons, 2012. Citado na página 44.
- COX, J.; ROSS, S.; RUBINSTEIN, M. Option pricing: A simplified approach. *Journal of financial Economics*, Elsevier, v. 7, n. 3, p. 229–263, 1979. Citado na página 29.
- CSILLÉRY, K.; BLUM, M.; GAGGIOTTI, O.; FRANÇOIS, O. Approximate bayesian computation (abc) in practice. *Trends in ecology & evolution*, Elsevier, v. 25, n. 7, p. 410–418, 2010. Citado na página 42.
- DEGROOT, M.; SCHERVISH, M. *Probability and statistics*. [S.l.]: Pearson Education, 2012. Citado na página 38.
- DURRETT, R. *Probability: theory and examples*. [S.l.]: Cambridge university press, 2010. Citado na página 17.
- EBERLEIN, E.; PRAUSE, K. The generalized hyperbolic model: financial derivatives and risk measures. In: SPRINGER. *Mathematical Finance—Bachelier Congress 2000*. [S.l.], 2002. p. 245–267. Citado na página 66.
- FLORENS-ZMIROU, D. Statistics on crossings of discretized diffusions and local time. *Stochastic Process. Appl.*, v. 39, p. 139–151, 1993. Citado na página 51.
- FOLKS, J.; CHHIKARA, R. The inverse gaussian distribution and its statistical application—a review. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 263–289, 1978. Citado na página 31.

- FRIEDLAENDER, A.; HALPIN, P.; QIAN, S.; LAWSON, G.; WIEBE, P.; THIELE, D.; READ, A. Whale distribution in relation to prey abundance and oceanographic processes in shelf waters of the western antarctic peninsula. *Marine Ecology Progress Series*, v. 317, p. 297–310, 2006. Citado na página 65.
- HALGREEN, C. Self-decomposability of the generalized inverse gaussian and hyperbolic distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, Springer, v. 47, n. 1, p. 13–17, 1979. Citado na página 31.
- HALMOS, P. R. *Lectures on ergodic theory*. [S.l.]: Courier Dover Publications, 2017. Citado na página 18.
- HARRISON, J.; SHEPP, L. On skew brownian motion. *The Annals of probability*, JSTOR, p. 309–313, 1981. Citado 2 vezes nas páginas 14 e 24.
- _____. On skew brownian motion. *Ann. Probab.*, v. 9, p. 309–313, 1981. Citado na página 23.
- HEYDE, C. A risky asset model with strong dependence through fractal activity time. *Journal of Applied Probability*, Cambridge University Press, v. 36, n. 4, p. 1234–1239, 1999. Citado 2 vezes nas páginas 14 e 31.
- HEYDE, C.; GAY, R. Fractals and contingent claims. *Preprint, Australian National University*, 2002. Citado na página 14.
- HEYDE, C.; GAY, R.; LIU, S. et al. Fractal scaling and black-scholes [a new view of long-range dependence in stock prices.]. *JASSA*, Financial Services Institute of Australasia, n. 1, p. 29, 2001. Citado na página 14.
- HEYDE, C.; LEONENKO, N. Student processes. *Advances in Applied Probability*, Cambridge University Press, v. 37, n. 2, p. 342–365, 2005. Citado 2 vezes nas páginas 14 e 31.
- HOWISON, S.; LAMPER, D. Trading volume in models of financial derivatives. *Applied Mathematical Finance*, Taylor & Francis, v. 8, n. 2, p. 119–135, 2001. Citado na página 32.
- HURST, S.; PLATEN, E. The marginal distributions of returns and volatility. *Lecture Notes-Monograph Series*, JSTOR, p. 301–314, 1997. Citado na página 14.
- JACOD, J. Rates of convergence to the local time of a diffusion. *Ann. Inst. H. Poincaré Probab. Statist.*, v. 34, p. 505–544, 1998. Citado 2 vezes nas páginas 51 e 56.
- JACOD, J.; PROTTER, P. *Probability essentials*. [S.l.]: Springer Science & Business Media, 2004. Citado 2 vezes nas páginas 17 e 19.
- JOYCE, P.; MARJORAM, P. et al. Approximately sufficient statistics and bayesian computation. *Statistical applications in genetics and molecular biology*, bepress, v. 7, n. 1, p. 26, 2008. Citado na página 44.
- KARLIS, D.; LILLESTÖL, J. Bayesian estimation of nig models via markov chain monte carlo methods. *Applied Stochastic Models in Business and Industry*, Wiley Online Library, v. 20, n. 4, p. 323–338, 2004. Citado na página 66.

- KERSS, A.; LEONENKO, N.; SIKORSKII, A. Risky asset models with tempered stable fractal activity time. *Stochastic Analysis and Applications*, Taylor & Francis, v. 32, n. 4, p. 642–663, 2014. Citado na página 31.
- KUTOYANTS, Y. *Parameter Estimation for Stochastic Processes*. [S.l.]: Heldermann, 1984. Citado na página 51.
- LEJAY, A. On the constructions of the skew brownian motion. *Probab. Surv.*, v. 3, p. 413–66, 2006. Citado 2 vezes nas páginas 51 e 53.
- LEJAY, A.; MORDECKI, E.; TORRES, S. Is a the brownian motion skew? *Scand. J. Statist.*, 2014. Citado 6 vezes nas páginas 14, 24, 51, 55, 56 e 57.
- LEUNG, E.; CHILVERS, B.; NAKAGAWA, S.; MOORE, A.; ROBERTSON, B. Sexual segregation in juvenile new zealand sea lion foraging ranges: implications for intraspecific competition, population dynamics and conservation. *PloS one*, v. 7, n. 9, p. e45389, 2012. Citado na página 65.
- LILLESTØL, J. *Bayesian estimation of NIG-parameters by Markov chain Monte Carlo methods*. [S.l.], 2000. Citado na página 66.
- LIPSTER, R.; SHIRYAEV, A. *Statistics of Random Processes*. [S.l.]: Springer, 2001. II. (Applications, II). Citado na página 51.
- LOSCHI, R.; SANTOS, C.; ARELLANO-VALLE, R. Combining bayesian procedures for testing. *risk*, v. 7, p. 16, 2010. Citado na página 40.
- MADRUGA, M. R.; ESTEVES, L.; WECHSLER, S. On the bayesianity of pereira-stern tests. *Test*, Springer, v. 10, n. 2, p. 291–299, 2001. Citado na página 41.
- MARCUS, M. *Xi-radial Processes and Random Fourier Series*. [S.l.]: American Mathematical Soc., 1987. Citado na página 30.
- MARIN, J.; PUDLO, P.; ROBERT, C.; RYDER, R. Approximate bayesian computational methods. *Statistics and Computing*, Springer, v. 22, n. 6, p. 1167–1180, 2012. Citado na página 43.
- MARJORAM, P.; MOLITOR, J.; PLAGNOL, V.; TAVARÉ, S. Markov chain monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 100, n. 26, p. 15324–15328, 2003. Citado 2 vezes nas páginas 14 e 43.
- MÖRTERS, P.; PERES, Y. *Brownian motion*. [S.l.]: Cambridge University Press, 2010. v. 30. Citado 2 vezes nas páginas 17 e 21.
- PEREIRA, C.; STERN, J.; WECHSLER, S. et al. Can a significance test be genuinely bayesian? *Bayesian Analysis*, International Society for Bayesian Analysis, v. 3, n. 1, p. 79–100, 2008. Citado na página 40.
- PRITCHARD, J.; SEIELSTAD, M.; PEREZ-LEZAUN, A.; FELDMAN, M. Population growth of human y chromosomes: a study of y chromosome microsatellites. *Molecular biology and evolution*, SMOE, v. 16, n. 12, p. 1791–1798, 1999. Citado na página 42.
- PROTASSOV, R. S. Em-based maximum likelihood parameter estimation for multivariate generalized hyperbolic distributions with fixed λ . *Statistics and Computing*, Springer, v. 14, n. 1, p. 67–77, 2004. Citado na página 66.

- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2013. Disponível em: <<http://www.R-project.org/>>. Citado na página 68.
- RATMANN, O.; ANDRIEU, C.; WIUF, C.; RICHARDSON, S. Model criticism based on likelihood-free inference, with an application to protein network evolution. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 106, n. 26, p. 10576–10581, 2009. Citado na página 14.
- RESNICK, S. *A probability path*. [S.l.]: Springer Science & Business Media, 2013. Citado na página 17.
- RIET-SAPRIZA, F.; COSTA, D.; FRANCO-TRECU, V.; MARÍN, Y.; AL., J. C. et. Foraging behavior of lactating south american sea lions (*otaria flavescens*) and spatial-temporal resource overlap with the uruguayan fisheries. *Deep Sea Res Part 2 Top Stud Oceanogr.*, v. 88–89, p. 106–119, 2013. Citado na página 63.
- RIFO, L.; GONZÁLEZ-LÓPEZ, V. Full bayesian analysis for a model of tail dependence. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 41, n. 22, p. 4107–4123, 2012. Citado na página 40.
- RIFO, L.; TORRES, S. Full bayesian analysis for a class of jump-diffusion models. *Communications in Statistics—Theory and Methods*, Taylor & Francis, v. 38, n. 8, p. 1262–1271, 2009. Citado na página 40.
- ROBSON, B.; GOEBEL, M.; BAKER, J.; REAM, R.; LOUGHLIN, T.; FRANCIS, R.; ANTONELIS, G.; COSTA, D. Separation of foraging habitat among breeding sites of a colonial marine predator, the northern fur seal (*callorhinus ursinus*). *Canadian Journal of Zoology*, v. 82, n. 1, p. 20–9, 2004. Citado na página 63.
- RODRIGUES, J. Bayesian analysis of zero-inflated distributions. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 32, n. 2, p. 281–289, 2003. Citado na página 40.
- RODRÍGUEZ, D.; DASSIS, M.; LEÓN, A. P. de; BARREIRO, C.; FARENGA, M.; AL. et. Foraging strategies of southern sea lion females in the la plata river estuary (argentina-uruguay). *Deep Sea Res Part 2 Top Stud Oceanogr.*, v. 88–89, p. 120–30, 2013. Citado na página 63.
- ROSINSKI, J. On a class of infinitely divisible processes represented as mixtures of gaussian processes. In: *Stable processes and related topics*. [S.l.]: Springer, 1991. p. 27–41. Citado na página 30.
- SATO, K. *Lévy processes and infinitely divisible distributions*. [S.l.]: Cambridge university press, 1999. Citado na página 26.
- SCHERVISH, M. J. *Theory of statistics*. [S.l.]: Springer Science & Business Media, 2012. Citado na página 38.
- SICHEL, H. On a distribution representing sentence-length in written prose. *Journal of the Royal Statistical Society. Series A (General)*, JSTOR, p. 25–34, 1974. Citado na página 31.

- SIGLER, M.; TOLLIT, D.; VOLLENWEIDER, J.; THEDINGA, J.; CSEPP, D.; WOMBLE, J.; TRITES, A. Steller sea lion foraging response to seasonal changes in prey availability. *Marine Ecology Progress Series*, v. 388, p. 243–61, 2009. Citado na página 65.
- SISSON, S.; FAN, Y.; TANAKA, M. Sequential monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 104, n. 6, p. 1760–1765, 2007. Citado 4 vezes nas páginas 14, 43, 44 e 47.
- SPIEGELHALTER, D.; BEST, N.; CARLIN, B.; LINDE, A. V. D. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 64, n. 4, p. 583–639, 2002. Citado na página 42.
- STEUTEL, F.; HARN, K. V. *Infinite divisibility of probability distributions on the real line*. [S.l.]: CRC Press, 2003. Citado na página 31.
- TAVARÉ, S.; BALDING, D.; GRIFFITHS, R.; DONNELLY, P. Inferring coalescence times from dna sequence data. *Genetics*, Genetics Soc America, v. 145, n. 2, p. 505–518, 1997. Citado na página 42.
- TORRES, S.; TUDOR, C. Donsker type theorem for the rosenblatt process and a binary market model. *Stochastic Analysis and Applications*, Taylor & Francis, v. 27, n. 3, p. 555–573, 2009. Citado na página 40.
- TURNER, B. M.; ZANDT, T. V. A tutorial on approximate bayesian computation. *Journal of Mathematical Psychology*, Elsevier, v. 56, n. 2, p. 69–85, 2012. Citado na página 44.
- VOSS, J. *An introduction to statistical computing: a simulation-based approach*. [S.l.]: John Wiley & Sons, 2013. Citado 2 vezes nas páginas 39 e 43.
- WAGNER, N. *Credit risk: Models, derivatives, and management*. [S.l.]: CRC press, 2008. Citado na página 14.
- WALSH, J. A diffusion with a discontinuous local time. *Astérisque*, v. 52, p. 37–45, 1978. Citado na página 52.

Anexos

ANEXO A – Códigos fonte

Os códigos-fonte a seguir são feitos na linguagem R,

A.0.1 Metropolis

```

1 #           Metropolis-Hastings
2 #
3 # prior t ~ Beta(ath,bth)
4 # prior v ~ invGamma(as,bs)
5 # data = data sample
6 # lp0,th0,v0 = initial state (posterior mode in grade =
    MHskew0(vdata) )
7 # e1, e2 - step sizes in MCMC
8 # N - sample size MCMC
9 MHskew1 <- function(ath,bth,as,bs,data,lp0,th0,v0,e1,e2,N){
10 skewMHth<-numeric(0);skewMHv<-numeric(0);skewMHlpost<-numeric
    (0);
11 skewMHth[1] <- th0;
12 skewMHv[1] <- v0;
13 skewMHlpost[1] <- lp0;
14 cont <- 1
15 for (i in 1:N){
16 R0 <- skewMHlpost[i];
17 th0i <- skewMHth[i]; v0i <- skewMHv[i];
18 epsth <- runif(1,-e1,e1);
19 thpot <- th0i+epsth
20 th1 <- ifelse(thpot< -1,-2-thpot,ifelse(thpot> 1,2-thpot,
    thpot));
21 epsv <- runif(1,-e2,e2);
22 vpot <- v0i+epsv;
23 v1 <- ifelse(vpot>0,vpot,v0i-epsv);
24 R1 <- lposteriorskew(th1,v1,ath,bth,as,bs,data)
25 u <- log10(runif(1));
26 alph <- R1-R0;
27 if (is.finite(alph) && alph >= u){
28 skewMHth[i+1] <- th1;
29 skewMHv[i+1] <- v1;

```

```

30 skewMHlpost[i+1] <- R1;
31 cont <- cont+1;
32 }
33 else{
34 skewMHth[i+1] <- th0i;
35 skewMHv[i+1] <- v0i;
36 skewMHlpost[i+1] <- R0;
37 }
38 }
39 psample <- cbind(skewMHth,skewMHv,skewMHlpost);
40 list(psample,cont/N)
41 }
42 #####
43 #           e-valor
44 #
45 evaluate <- function(th,v,lp,h0){
46 d0 <- abs(th-h0)
47 q5 <- quantile(d0,probs=.05,na.rm=TRUE)
48 mh0 <- max(lp[which(d0<q5)],0,na.rm=T)
49 1-length(which(lp>mh0))/length(th)
50 }

```

A.0.2 ABC

```

1 #           ABC-PRC
2 #
3 abc.pmc <- function(npart,niter,mu,sigma_nor,parexp,y,eps,sig
, the){
4 if(length(eps) != niter)stop("eps errado")
5 if(niter < 20)nscore <- niter
6 else nscore <- 20
7 parameterEstimating <- array(dim=c(nscore,npart,2))
8 p.historyt <- matrix(nrow=nscore,ncol=npart)
9 p.history <- matrix(nrow=nscore,ncol=npart)
10 pt <- round(seq(1,niter,length=nscore))
11 nsamp <- length(y)
12 yrho <- rho(y)
13 v1 <- rnorm(npart,mu,sigma_nor)
14 wtvec <- rep(1,npart)
15 k <- 1

```

```

16 for(j in 1:niter){{
17   vx <- numeric(0)
18   while(T){
19     vv <- sample(v1,100*npart,replace=T,prob=wtvec)#,prob=wtvec
20     ssvec <- lapply(lapply(vv,simtheta,sigma=sig),rho)
21     ind <- sqrt((as.numeric(ssvec) - yrho)^2) <= eps[j]^2
22     if(sum(ind) == 0)next;
23     vx <- c(vx,vv[ind])
24     if(length(vx) >= npart){
25       vx <- vx[1:npart]
26       break;
27     }
28   }
29   if(pt[k] == j){
30     p.historyt[k,] <- vx;
31     k <- k+1
32   }
33   aux=wtvec
34   dem=numeric(0)
35   v1 <- vx + rnorm(npart,0,sqrt(.1))
36   for(jj in 1: npart){
37     for(l in 1: npart){
38       dem[l]=aux[l]*dnorm(vx[l],v1[l],sqrt(.1))
39     }
40     wtvec[jj] <- dnorm(v1[jj],mu,sigma_nor)/sum(dem)
41   }
42 }
43 parameterEstimating[, ,1] <- t(p.historyt)}
44 ysd <- var(y)
45 v1 <- rexp(npart,parexp)
46 wtvec <- rep(1,npart)
47 k <- 1
48 for(j in 1:niter){{
49   vx <- numeric(0)
50   while(T){
51     vv <- sample(v1,100*npart,replace=T,prob=wtvec)#,prob=wtvec
52     ssvec <- lapply(lapply(vv,simsigma,theta=the),var)
53     ind <- sqrt((as.numeric(ssvec) - ysd)^2) <= eps[j]^2
54     if(sum(ind) == 0)next;

```

```
55 vx <- c(vx,vv[ind])
56 if(length(vx) >= npart){
57 vx <- vx[1:npart]
58 break;
59 }
60 }
61 if(pt[k] == j){
62 p.history[k,] <- vx;
63 k <- k+1
64 }
65
66 aux=wvec
67 dem=numeric(0)
68 v1 <- vx + rnorm(npart,0,sqrt(.1))
69 v1 <- abs(v1)
70 for(jj in 1: npart){
71 for(l in 1: npart){
72 dem[l]=aux[l]*dnorm(vx[l],v1[l],sqrt(.1))
73 }
74 wvec[jj] <- dexp(v1[jj],parexp)/sum(dem)
75 }
76 }
77 parameterEstimating[, ,2] <-t(p.history)}
78 parameterEstimating
79 }
```